

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 101.709 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-13
Perfect score: 73
Sequence: 1 IEGPTLRQWLARA 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	14	2	AAW09463
2	73	100.0	14	2	AAW09468
3	73	100.0	14	2	AAW33030
4	73	100.0	14	2	AAW33034
5	73	100.0	14	2	AAW36774
6	73	100.0	14	2	ADI24843
7	73	100.0	14	3	AAAY96515
8	73	100.0	14	3	AAAB16962
9	73	100.0	14	4	AAU25827
10	73	100.0	14	4	AAU26004
11	73	100.0	14	5	ABBB72853
12	73	100.0	14	5	ABBP51669
13	73	100.0	14	5	AAAE18011
14	73	100.0	14	6	ABGG71747
15	73	100.0	14	7	ABRB62907
16	73	100.0	14	7	ADCC33697
17	73	100.0	14	7	ADN59652
18	73	100.0	14	8	ADL27293
19	73	100.0	14	8	ADM72483
20	73	100.0	14	8	ADQ16584
21	73	100.0	14	8	ADT92482
22	73	100.0	14	9	ADU70210
23	73	100.0	14	9	ADU75982
24	73	100.0	14	9	ADV44319

25	73	100.0	14	9	AEB12792	Aeb12792	TPO mimet
26	73	100.0	15	2	AAW35416	Aaw35416	Thrombopo
27	73	100.0	15	2	AAW36776	Aaw36776	Thrombopo
28	73	100.0	15	2	AAW66712	Aaw66712	Peptide c
29	73	100.0	15	3	AAB20684	Aab20684	Thrombocy
30	73	100.0	15	4	AAU25996	Aau25996	Human thr
31	73	100.0	15	4	AAU25831	Aau25831	Human thr
32	73	100.0	15	5	ABP51670	Abp51670	Thrombopo
33	73	100.0	15	7	ABR62908	Abr62908	Thrombopo
34	73	100.0	15	8	ADM72485	Adm72485	TPO mimet
35	73	100.0	15	8	ADM72479	Adm72479	TPO mimet
36	73	100.0	15	8	ADM72478	Adm72478	TPO mimet
37	73	100.0	15	8	ADM72533	Adm72533	TPO mimet
38	73	100.0	15	8	ADM72522	Adm72522	TPO mimet
39	73	100.0	15	8	ADM72523	Adm72523	TPO mimet
40	73	100.0	15	8	ADM72482	Adm72482	TPO mimet
41	73	100.0	15	8	ADQ16585	Adq16585	TPO mimet
42	73	100.0	15	8	ADT92483	Adt92483	Modified
43	73	100.0	15	9	ADU70209	Adu70209	Thrombopo
44	73	100.0	15	9	ADU75981	Adu75981	Peptide-b
45	73	100.0	15	9	ADV44320	Adv44320	Agonist T
46	73	100.0	15	9	AEB12793	Aeb12793	TPO mimet
47	73	100.0	16	2	AAW19534	Aaw19534	Thrombopo
48	73	100.0	16	2	AAW33035	Aaw33035	Thrombopo
49	73	100.0	16	2	AAW36775	Aaw36775	Thrombopo
50	73	100.0	16	2	AAW36771	Aaw36771	Thrombopo
51	73	100.0	16	2	AAW66709	Aaw66709	Peptide c
52	73	100.0	16	2	AAW66713	Aaw66713	Peptide c
53	73	100.0	16	2	AAW66733	Aaw66733	Peptide c
54	73	100.0	16	2	AAW66716	Aaw66716	Peptide c
55	73	100.0	16	4	AAU26005	Aau26005	Human thr
56	73	100.0	16	4	AAU26043	Aau26043	Human thr
57	73	100.0	16	4	AAU25832	Aau25832	Human thr
58	73	100.0	16	8	ADM72532	Adm72532	TPO mimet
59	73	100.0	16	8	ADM72484	Adm72484	TPO mimet
60	73	100.0	18	3	AAB16957	Aab16957	PEGylated
61	73	100.0	18	3	AAB16956	Aab16956	PEGylated
62	73	100.0	18	5	ABP51687	Abp51687	TPO mimet
63	73	100.0	18	5	ABP51689	Abp51689	TPO mimet
64	73	100.0	18	5	ABP51688	Abp51688	TPO mimet
65	73	100.0	18	5	ABP51677	Abp51677	TPO mimet
66	73	100.0	18	5	ABP51686	Abp51686	TPO mimet
67	73	100.0	18	5	ABP51674	Abp51674	TPO mimet
68	73	100.0	18	5	ABP51693	Abp51693	TPO mimet
69	73	100.0	18	5	ABP51684	Abp51684	TPO mimet
70	73	100.0	18	5	ABP51683	Abp51683	TPO mimet
71	73	100.0	18	5	ABP51685	Abp51685	TPO mimet
72	73	100.0	18	5	ABP51691	Abp51691	TPO mimet
73	73	100.0	18	5	ABP51673	Abp51673	TPO mimet
74	73	100.0	18	5	ABP51690	Abp51690	TPO mimet
75	73	100.0	18	5	ABP51675	Abp51675	TPO mimet
76	73	100.0	18	5	ABP51692	Abp51692	TPO mimet
77	73	100.0	18	7	ADN59812	Adn59812	Thrombopo
78	73	100.0	18	8	ADQ16611	Adq16611	TPO mimet
79	73	100.0	18	8	ADQ16619	Adq16619	TPO mimet
80	73	100.0	18	8	ADQ16621	Adq16621	TPO mimet
81	73	100.0	18	8	ADQ16641	Adq16641	TPO mimet
82	73	100.0	18	8	ADQ16646	Adq16646	TPO mimet
83	73	100.0	18	8	ADQ16607	Adq16607	TPO mimet
84	73	100.0	18	8	ADQ16615	Adq16615	TPO mimet
85	73	100.0	18	8	ADQ16627	Adq16627	TPO mimet
86	73	100.0	18	8	ADQ16625	Adq16625	TPO mimet
87	73	100.0	18	8	ADQ16617	Adq16617	TPO mimet
88	73	100.0	18	8	ADQ16629	Adq16629	TPO mimet
89	73	100.0	18	8	ADQ16613	Adq16613	TPO mimet
90	73	100.0	18	8	ADQ16623	Adq16623	TPO mimet
91	73	100.0	18	8	ADQ16605	Adq16605	TPO mimet
92	73	100.0	18	8	ADQ16609	Adq16609	TPO mimet
93	73	100.0	18	9	ADV44355	Adv44355	Agonist T
94	73	100.0	18	9	ADV44345	Adv44345	Agonist T
95	73	100.0	18	9	ADV44351	Adv44351	Agonist T
96	73	100.0	18	9	ADV44357	Adv44357	Agonist T
97	73	100.0	18	9	ADV44349	Adv44349	Agonist T

98	73	100.0	18	9	ADV44343	Adv44343 Agonist T	171	73	100.0	29	3	AAB16975	Aab16975 TPO-mimet
99	73	100.0	18	9	ADV44347	Adv44347 Agonist T	172	73	100.0	29	3	AAB16976	Aab16976 TPO-mimet
100	73	100.0	18	9	ADV44384	Adv44384 Agonist T	173	73	100.0	29	3	AAB17286	Aab17286 TPO-mimet
101	73	100.0	18	9	ADV44353	Adv44353 Agonist T	174	73	100.0	29	3	AAB16970	Aab16970 TPO-mimet
102	73	100.0	18	9	ADV44363	Adv44363 Agonist T	175	73	100.0	29	5	ABB72862	Abb72862 TPO mimet
103	73	100.0	18	9	ADV44361	Adv44361 Agonist T	176	73	100.0	29	5	ABB72861	Abb72861 TPO mimet
104	73	100.0	18	9	ADV44379	Adv44379 Agonist T	177	73	100.0	29	5	ABB72856	Abb72856 TPO mimet
105	73	100.0	18	9	ADV44344	Adv44344 Agonist T	178	73	100.0	29	7	ADJ73011	Adj73011 TPO mimet
106	73	100.0	18	9	ADV44359	Adv44359 Agonist T	179	73	100.0	29	7	ADJ73006	Adj73006 TPO mimet
107	73	100.0	18	9	ADV44365	Adv44365 Agonist T	180	73	100.0	29	8	ADJ52646	Adj52646 CH1 delet
108	73	100.0	18	9	ADV44367	Adv44367 Agonist T	181	73	100.0	29	8	ADJ52641	Adj52641 CH1 delet
109	73	100.0	18	9	ADV44468	Adv44468 Agonist T	182	73	100.0	29	8	ADJ51602	Adj51602 CH1 delet
110	73	100.0	18	9	AEB12826	Aeb12826 TPO mimet	183	73	100.0	29	8	ADJ51607	Adj51607 CH1 delet
111	73	100.0	18	9	AEB12820	Aeb12820 TPO mimet	184	73	100.0	30	3	AAB17287	Aab17287 TPO-mimet
112	73	100.0	18	9	AEB12840	Aeb12840 TPO mimet	185	73	100.0	31	3	AAB17288	Aab17288 TPO-mimet
113	73	100.0	18	9	AEB12857	Aeb12857 Antibody	186	73	100.0	31	3	AAB16974	Aab16974 TPO-mimet
114	73	100.0	18	9	AEB12834	Aeb12834 TPO mimet	187	73	100.0	31	3	AAB16973	Aab16973 TPO-mimet
115	73	100.0	18	9	AEB12830	Aeb12830 TPO mimet	188	73	100.0	31	5	ABB72860	Abb72860 TPO mimet
116	73	100.0	18	9	AEB12939	Aeb12939 TPO mimet	189	73	100.0	31	5	ABB72859	Abb72859 TPO mimet
117	73	100.0	18	9	AEB12832	Aeb12832 TPO mimet	190	73	100.0	31	7	ADJ73009	Adj73009 TPO mimet
118	73	100.0	18	9	AEB12852	Aeb12852 TPO mimet	191	73	100.0	31	7	ADJ73010	Adj73010 TPO mimet
119	73	100.0	18	9	AEB12816	Aeb12816 TPO mimet	192	73	100.0	31	8	ADJ52644	Adj52644 CH1 delet
120	73	100.0	18	9	AEB12818	Aeb12818 TPO mimet	193	73	100.0	31	8	ADJ52645	Adj52645 CH1 delet
121	73	100.0	18	9	AEB12828	Aeb12828 TPO mimet	194	73	100.0	31	8	ADJ51606	Adj51606 CH1 delet
122	73	100.0	18	9	AEB12836	Aeb12836 TPO mimet	195	73	100.0	31	8	ADJ51605	Adj51605 CH1 delet
123	73	100.0	18	9	AEB12822	Aeb12822 TPO mimet	196	73	100.0	32	3	AAY96520	Aay96520 Thrombopo
124	73	100.0	18	9	AEB12824	Aeb12824 TPO mimet	197	73	100.0	32	3	AAB17289	Aab17289 TPO-mimet
125	73	100.0	18	9	AEB12838	Aeb12838 TPO mimet	198	73	100.0	32	3	AAB17297	Aab17297 TPO-mimet
126	73	100.0	19	5	ABB73391	Abb73391 TPO-mimet	199	73	100.0	33	3	AAB17290	Aab17290 TPO-mimet
127	73	100.0	19	5	ABB73390	Abb73390 TPO-mimet	200	73	100.0	34	3	AAY96527	Aay96527 Thrombopo
128	73	100.0	20	3	AAB18003	Aab18003 Fc-TMP pe	201	73	100.0	34	3	AAB17291	Aab17291 TPO-mimet
129	73	100.0	20	3	AAB17929	Aab17929 TPO-mimet	202	73	100.0	35	3	AAB17292	Aab17292 TPO-mimet
130	73	100.0	20	5	ABB73403	Abb73403 TPO mimet	203	73	100.0	36	3	AAY96525	Aay96525 Thrombopo
131	73	100.0	21	7	ADN59687	Adn59687 Thrombopo	204	73	100.0	36	3	AAY96523	Aay96523 Thrombopo
132	73	100.0	22	7	ADN59819	Adn59819 TMP pepti	205	73	100.0	36	3	AAY96524	Aay96524 Thrombopo
133	73	100.0	22	8	ADQ16714	Adq16714 Immunoglo	206	73	100.0	36	3	AAY96526	Aay96526 Thrombopo
134	73	100.0	22	8	ADQ16713	Adq16713 Immunoglo	207	73	100.0	36	3	AAB17307	Aab17307 TPO-mimet
135	73	100.0	22	8	ADQ16709	Adq16709 Immunoglo	208	73	100.0	36	3	AAB17293	Aab17293 TPO-mimet
136	73	100.0	22	8	ADQ16706	Adq16706 Immunoglo	209	73	100.0	36	3	AAB17303	Aab17303 TPO-mimet
137	73	100.0	22	8	ADQ16699	Adq16699 TPO mimet	210	73	100.0	36	3	AAB16963	Aab16963 TPO-mimet
138	73	100.0	22	8	ADQ16712	Adq16712 Immunoglo	211	73	100.0	36	3	AAB17301	Aab17301 TPO-mimet
139	73	100.0	22	8	ADQ16707	Adq16707 Immunoglo	212	73	100.0	36	3	AAB17306	Aab17306 TPO-mimet
140	73	100.0	22	8	ADQ16711	Adq16711 Immunoglo	213	73	100.0	36	5	ABB72403	Abb72403 TPO-mimet
141	73	100.0	22	8	ADQ16708	Adq16708 Immunoglo	214	73	100.0	37	3	AAB17294	Aab17294 TPO-mimet
142	73	100.0	22	8	ADQ16710	Adq16710 Immunoglo	215	73	100.0	38	3	AAB17295	Aab17295 TPO-mimet
143	73	100.0	22	9	ADV44435	Adv44435 Modified	216	73	100.0	39	3	AAB17304	Aab17304 TPO-mimet
144	73	100.0	22	9	ADV44449	Adv44449 Anti-teta	217	73	100.0	39	3	AAB17305	Aab17305 TPO-mimet
145	73	100.0	22	9	ADV44443	Adv44443 Anti-teta	218	73	100.0	40	3	AAB17302	Aab17302 TPO-mimet
146	73	100.0	22	9	ADV44444	Adv44444 Anti-teta	219	73	100.0	41	3	AAY96528	Aay96528 Thrombopo
147	73	100.0	22	9	ADV44448	Adv44448 Anti-teta	220	73	100.0	41	5	ABB73389	Abb73389 TPO-mimet
148	73	100.0	22	9	ADV44442	Adv44442 Anti-teta	221	73	100.0	41	5	ABB73388	Abb73388 TPO-mimet
149	73	100.0	22	9	ADV44447	Adv44447 Anti-teta	222	73	100.0	42	3	AAY96530	Aay96530 Thrombopo
150	73	100.0	22	9	ADV44445	Adv44445 Anti-teta	223	73	100.0	42	3	AAB17296	Aab17296 TPO-mimet
151	73	100.0	22	9	ADV44450	Adv44450 Anti-teta	224	73	100.0	42	3	AAB17308	Aab17308 Synthetic
152	73	100.0	22	9	ADV44446	Adv44446 Anti-teta	225	73	100.0	42	3	AAB17282	Aab17282 TPO-mimet
153	73	100.0	22	9	AEB12918	Aeb12918 TPO mimet	226	73	100.0	42	3	AAB17281	Aab17281 TPO-mimet
154	73	100.0	22	9	AEB12917	Aeb12917 TPO mimet	227	73	100.0	42	5	ABB73404	Abb73404 TMP-TMP g
155	73	100.0	22	9	AEB12919	Aeb12919 TPO mimet	228	73	100.0	60	3	AAB17311	Aab17311 Synthetic
156	73	100.0	22	9	AEB12986	Aeb12986 TT antibo	229	73	100.0	60	5	ABB73405	Abb73405 TMP-TMP g
157	73	100.0	22	9	AEB12920	Aeb12920 TPO mimet	230	73	100.0	122	9	ADV44474	Adv44474 Anti-teta
158	73	100.0	22	9	AEB12916	Aeb12916 TPO mimet	231	73	100.0	122	9	AEB12946	Aeb12946 Antibody
159	73	100.0	22	9	AEB12914	Aeb12914 TPO mimet	232	73	100.0	128	8	ADQ16705	Adq16705 Modified
160	73	100.0	22	9	AEB12921	Aeb12921 TPO mimet	233	73	100.0	128	9	ADV44463	Adv44463 Anti-teta
161	73	100.0	22	9	AEB12907	Aeb12907 TT antibo	234	73	100.0	128	9	ADV44467	Adv44467 Anti-teta
162	73	100.0	22	9	AEB12915	Aeb12915 TPO mimet	235	73	100.0	128	9	ADV44465	Adv44465 Anti-teta
163	73	100.0	28	3	AAB17285	Aab17285 TPO-mimet	236	73	100.0	128	9	ADV44461	Adv44461 pAX116 va
164	73	100.0	28	5	ABP51682	Abp51682 TPO mimet	237	73	100.0	128	9	ADV44464	Adv44464 Anti-teta
165	73	100.0	28	7	ADJ73013	Adj73013 TPO mimet	238	73	100.0	128	9	AEB12934	Aeb12934 Antibody
166	73	100.0	28	8	ADJ52648	Adj52648 CH1 delet	239	73	100.0	128	9	AEB12935	Aeb12935 Antibody
167	73	100.0	28	8	ADJ51609	Adj51609 CH1 delet	240	73	100.0	128	9	AEB12913	Aeb12913 Antibody
168	73	100.0	28	8	ADQ16636	Adq16636 Tetanus t	241	73	100.0	128	9	AEB12936	Aeb12936 Antibody
169	73	100.0	28	9	ADV44374	Adv44374 Modified	242	73	100.0	128	9	AEB12937	Aeb12937 Antibody
170	73	100.0	28	9	AEB12847	Aeb12847 Antibody	243	73	100.0	128	9	AEB12937	Aeb12937 Antibody

244	73	100.0	128	9	9	AEB12938	Aeb12938	Antibody	317	67	91.8	15	8	ADM72496	Adm72496	TPO mimet
245	73	100.0	129	6	6	ABG71751	Abg71751	Antibody	318	67	91.8	15	8	ADM72494	Adm72494	TPO mimet
246	73	100.0	131	6	6	ABG71753	Abg71753	Antibody	319	66	90.4	13	4	AAU26008	Aau26008	Human thr
247	73	100.0	132	9	6	ADV44473	Adv44473	Anti-teta	320	65	89.0	12	2	AAW36787	Aaw36787	Thrombopo
248	73	100.0	133	6	6	ABG71752	Abg71752	Antibody	321	65	89.0	13	4	AAU26012	Aau26012	Human thr
249	73	100.0	135	6	6	ABG71749	Abg71749	Antibody	322	65	89.0	14	2	AAW36788	Aaw36788	Thrombopo
250	73	100.0	143	6	6	ABG71750	Abg71750	Antibody	323	65	89.0	14	4	AAU26013	Aau26013	Human thr
251	73	100.0	144	6	6	ABG71748	Abg71748	Antibody	324	65	89.0	15	2	AAW66717	Aaw66717	Peptide c
252	73	100.0	150	9	9	AEB12945	Aeb12945	Antibody	325	64	87.7	12	2	AAW36781	Aaw36781	Thrombopo
253	73	100.0	225	8	8	ADQ16704	Adq16704	Modified	326	64	87.7	14	2	AAW36782	Aaw36782	Thrombopo
254	73	100.0	234	9	9	AEB12912	Aeb12912	Antibody	327	64	87.7	14	4	AAU26009	Aau26009	Human thr
255	73	100.0	247	3	3	AAB16958	Aab16958	FC-TMP pr	328	64	87.7	18	7	ADN59663	Adn59663	Thrombopo
256	73	100.0	247	3	3	AAB16961	Aab16961	TMP-Fc pr	329	64	87.7	22	7	ADN59830	Adn59830	TMP pepti
257	73	100.0	247	5	5	ABB73411	Abb73411	FC-TPO mi	330	64	87.7	25	7	ADN59708	Adn59708	Thrombopo
258	73	100.0	247	5	5	ABB73414	Abb73414	TMP-Fc am	331	64	87.7	43	7	ADN59759	Adn59759	Peptide-v
259	73	100.0	249	9	9	ADV44440	Adv44440	pAX116 va	332	62	84.9	12	4	AAU26014	Aau26014	Human thr
260	73	100.0	268	3	3	AAB16959	Aab16959	FC-TMP-TM	333	61	83.6	14	4	AAU26037	Aau26037	Human thr
261	73	100.0	268	5	5	ABB73412	Abb73412	FC-TMP-TM	334	61	83.6	14	8	ADM72526	Adm72526	TPO mimet
262	73	100.0	269	3	3	AAV96531	Aay96531	Human IGG	335	61	83.6	14	8	ADM72499	Adm72499	TPO mimet
263	73	100.0	269	3	3	AAB16960	Aab16960	TMP-TMP-F	336	61	83.6	14	8	ADM72527	Adm72527	TPO mimet
264	73	100.0	269	5	5	ABB73413	Abb73413	TMP-TMP-F	337	61	83.6	15	2	AAW66731	Aaw66731	Peptide c
265	73	100.0	282	9	9	AEB12930	Aeb12930	Antibody	338	61	83.6	15	4	AAU26024	Aau26024	Human thr
266	73	100.0	459	9	9	ADV44459	Adv44459	Anti-teta	339	61	83.6	15	4	AAU26038	Aau26038	Human thr
267	73	100.0	472	5	5	ABP51695	Abp51695	SG1.1-TPO	340	61	83.6	15	8	ADM72498	Adm72498	TPO mimet
268	73	100.0	472	8	8	ADQ16647	Adq16647	Immunoglo	341	60	82.2	13	2	AAW36792	Aaw36792	Thrombopo
269	73	100.0	472	9	9	ADV44385	Adv44385	SG1.1 hea	342	60	82.2	13	4	AAU26015	Aau26015	Human thr
270	73	100.0	472	9	9	AEB12858	Aeb12858	Antibody	343	60	82.2	14	8	ADM72475	Adm72475	TPO mimet
271	70	95.9	14	3	3	AAB16969	Aab16969	TPO-mimet	344	60	82.2	14	9	ADY64334	Ady64334	Thrombopo
272	70	95.9	14	3	3	AAB16968	Aab16968	TPO-mimet	345	60	82.2	15	9	ADM72477	Adm72477	TPO mimet
273	70	95.9	14	5	5	ABB72854	Abb72854	TPO mimet	346	60	82.2	15	8	ADY64335	Ady64335	Thrombopo
274	70	95.9	14	5	5	ABB72855	Abb72855	TPO mimet	347	60	82.2	16	2	AAW66711	Aaw66711	Peptide c
275	70	95.9	14	7	7	ADJ73005	Adj73005	TPO mimet	348	60	82.2	16	8	ADM72476	Adm72476	TPO mimet
276	70	95.9	14	7	7	ADJ73004	Adj73004	TPO mimet	349	60	82.2	19	2	AAW09491	Aaw09491	Thrombopo
277	70	95.9	14	8	8	ADJ52639	Adj52639	CH1 delet	350	60	82.2	19	2	AAW09493	Aaw09493	Thrombopo
278	70	95.9	14	8	8	ADJ52640	Adj52640	CH1 delet	351	60	82.2	19	2	AAW36644	Aaw36644	Thrombopo
279	70	95.9	14	8	8	ADJ51601	Adj51601	CH1 delet	352	60	82.2	19	2	AAW35418	Aaw35418	Thrombopo
280	70	95.9	14	8	8	ADJ51600	Adj51600	CH1 delet	353	60	82.2	19	2	AAW36642	Aaw36642	Thrombopo
281	70	95.9	15	2	2	AAW66718	Aaw66718	Peptide c	354	60	82.2	19	4	AAU25861	Aau25861	Human thr
282	70	95.9	29	3	3	AAB16971	Aab16971	TPO-mimet	355	60	82.2	19	4	AAU25863	Aau25863	Human thr
283	70	95.9	29	5	5	ABB72857	Abb72857	TPO mimet	356	60	82.2	19	4	AAU25998	Aau25998	Human thr
284	70	95.9	29	7	7	ADJ73007	Adj73007	TPO mimet	357	60	82.2	28	8	ADJ52647	Adj52647	CH1 delet
285	70	95.9	29	8	8	ADJ52642	Adj52642	CH1 delet	358	60	82.2	29	3	AAB16972	Aab16972	TPO-mimet
286	70	95.9	29	8	8	ADJ51603	Adj51603	CH1 delet	359	60	82.2	29	5	ABB72858	Abb72858	TPO mimet
287	69	94.5	13	2	2	AAW36779	Aaw36779	Thrombopo	360	60	82.2	29	7	ADJ73012	Adj73012	TPO mimet
288	69	94.5	13	7	7	ADJ73003	Adj73003	TPO mimet	361	60	82.2	29	7	ADJ73008	Adj73008	TPO mimet
289	69	94.5	13	8	8	ADJ52638	Adj52638	CH1 delet	362	60	82.2	29	8	ADJ52643	Adj52643	CH1 delet
290	69	94.5	13	8	8	ADJ51599	Adj51599	CH1 delet	363	60	82.2	29	8	ADJ51604	Adj51604	CH1 delet
291	69	94.5	14	4	4	AAU26006	Aau26006	Human thr	364	60	82.2	29	8	ADJ51608	Adj51608	CH1 delet
292	69	94.5	14	4	4	AAU26010	Aau26010	Human thr	365	60	82.2	36	3	AAV96521	Aay96521	Cyclic or
293	69	94.5	14	8	8	ADM72503	Adm72503	TPO mimet	366	60	82.2	36	3	AAB17298	Aab17298	TPO-mimet
294	69	94.5	14	8	8	ADM72487	Adm72487	TPO mimet	367	60	82.2	36	3	AAB17299	Aab17299	TPO-mimet
295	69	94.5	14	9	9	ADU70206	Adu70206	Thrombopo	368	59	80.8	14	8	ADM72509	Adm72509	TPO mimet
296	69	94.5	14	9	9	ADU75979	Adu75979	Thrombopo	369	59	80.8	15	2	AAW66723	Aaw66723	Peptide c
297	69	94.5	15	2	2	AAW36784	Aaw36784	Thrombopo	370	59	80.8	15	4	AAU26028	Aau26028	Human thr
298	69	94.5	15	2	2	AAW36780	Aaw36780	Thrombopo	371	59	80.8	15	8	ADM72508	Adm72508	TPO mimet
299	69	94.5	15	2	2	AAW66721	Aaw66721	Peptide c	372	59	80.8	18	2	AAW09460	Aaw09460	Thrombopo
300	69	94.5	15	2	2	AAW66721	Aaw66721	Peptide c	373	59	80.8	18	2	AAW09498	Aaw09498	Thrombopo
301	69	94.5	15	4	4	AAU26026	Aau26026	Human thr	374	59	80.8	18	2	AAW36649	Aaw36649	Thrombopo
302	69	94.5	15	4	4	AAU26011	Aau26011	Human thr	375	59	80.8	18	2	AAW33027	Aaw33027	Thrombopo
303	69	94.5	15	4	4	AAU26020	Aau26020	Human thr	376	59	80.8	18	2	AAW36652	Aaw36652	Thrombopo
304	69	94.5	15	4	4	AAU26007	Aau26007	Human thr	377	59	80.8	18	3	AAB17026	Aab17026	TPO-mimet
305	69	94.5	15	8	8	ADM72502	Adm72502	TPO mimet	378	59	80.8	18	4	AAU25868	Aau25868	Human thr
306	69	94.5	15	8	8	ADM72492	Adm72492	TPO mimet	379	59	80.8	18	4	AAU25824	Aau25824	Human thr
307	69	94.5	15	8	8	ADM72490	Adm72490	TPO mimet	380	59	80.8	18	4	AAU25871	Aau25871	Human thr
308	69	94.5	15	8	8	ADM72486	Adm72486	TPO mimet	381	59	80.8	18	5	ABB72912	Abb72912	TPO mimet
309	69	94.5	15	8	8	ADM72491	Adm72491	TPO mimet	382	59	80.8	18	7	ADJ73064	Adj73064	TPO mimet
310	69	94.5	15	8	8	ADM72493	Adm72493	TPO mimet	383	59	80.8	18	8	ADJ52699	Adj52699	CH1 delet
311	69	94.5	16	4	4	AAU26021	Aau26021	Human thr	384	59	80.8	18	8	ADJ51660	Adj51660	CH1 delet
312	67	91.8	14	8	8	ADM72495	Adm72495	TPO mimet	385	59	80.8	36	3	AAV96522	Aay96522	Linear th
313	67	91.8	14	8	8	ADM72497	Adm72497	TPO mimet	386	59	80.8	36	3	AAB17300	Aab17300	TPO-mimet
314	67	91.8	15	2	2	AAW66719	Aaw66719	Peptide c	387	58	79.5	13	2	AAW36783	Aaw36783	Thrombopo
315	67	91.8	15	4	4	AAU26022	Aau26022	Human thr	388	58	79.5	13	4	AAU26035	Aau26035	Human thr
316	67	91.8	15	4	4	AAU26023	Aau26023	Human thr	389	58	79.5	13	8	ADM72524	Adm72524	TPO mimet

390 58 79.5 14 2 AAW36773 Aaw36773 Thrombopo 463
391 57 78.1 13 4 AAU26018 Aau26018 Human thr 464
392 57 78.1 13 8 ADM72488 Adm72488 TPO mimet 465
393 57 78.1 18 2 AAW09499 Aaw09499 Thrombopo 466
394 57 78.1 18 2 AAW09459 Aaw09459 Thrombopo 467
395 57 78.1 18 2 AAW36650 Aaw36650 Thrombopo 468
396 57 78.1 18 2 AAW33026 Aaw33026 Thrombopo 469
397 57 78.1 18 3 AAB17024 Aab17024 TPO-mimet 470
398 57 78.1 18 4 AAU25869 Aau25869 Human thr 471
399 57 78.1 18 4 AAU25823 Aau25823 Human thr 472
400 57 78.1 18 5 ABB72910 Abb72910 TPO mimet 473
401 57 78.1 18 7 ADJ73062 Adj73062 TPO mimet 474
402 57 78.1 18 8 ADJ52697 Adj52697 CH1 delet 475
403 57 78.1 18 8 ADJ51658 Adj51658 CH1 delet 476
404 56 76.7 10 3 AAB17006 Aab17006 TPO-mimet 477
405 56 76.7 10 5 ABB72892 Abb72892 TPO mimet 478
406 56 76.7 10 7 ADJ73043 Adj73043 TPO mimet 479
407 56 76.7 10 8 ADJ52678 Adj52678 CH1 delet 480
408 56 76.7 10 8 ADJ51639 Adj51639 CH1 delet 481
409 56 76.7 13 8 ADM72525 Adm72525 TPO mimet 482
410 56 76.7 13 8 ADM72489 Adm72489 TPO mimet 483
411 56 76.7 14 2 AAW66715 Aaw66715 Peptide c 484
412 56 76.7 14 2 AAW66730 Aaw66730 Peptide c 485
413 56 76.7 14 4 AAU26019 Aau26019 Human thr 486
414 56 76.7 14 4 AAU26036 Aau26036 Human thr 487
415 56 76.7 14 8 ADM72507 Adm72507 TPO mimet 488
416 56 76.7 14 8 ADM72505 Adm72505 TPO mimet 489
417 56 76.7 14 8 ADM72501 Adm72501 TPO mimet 490
418 56 76.7 14 9 ADU70205 Adu70205 Thrombopo 491
419 56 76.7 14 9 ADU75978 Adu75978 Thrombopo 492
420 56 76.7 15 2 AAW66722 Aaw66722 Peptide c 493
421 56 76.7 15 2 AAW66720 Aaw66720 Peptide c 494
422 56 76.7 15 4 AAU25833 Aau25833 Human thr 495
423 56 76.7 15 4 AAU26027 Aau26027 Human thr 496
424 56 76.7 15 4 AAU26025 Aau26025 Human thr 497
425 56 76.7 15 8 ADM72506 Adm72506 TPO mimet 498
426 56 76.7 15 8 ADM72500 Adm72500 TPO mimet 499
427 56 76.7 15 8 ADM72504 Adm72504 TPO mimet 500
428 56 76.7 18 7 ADN59664 Adn59664 Thrombopo 501
429 56 76.7 18 7 ADN59660 Adn59660 Thrombopo 502
430 56 76.7 19 2 AAW09494 Aaw09494 Thrombopo 503
431 56 76.7 19 2 AAW09461 Aaw09461 Thrombopo 504
432 56 76.7 19 2 AAW33028 Aaw33028 Thrombopo 505
433 56 76.7 19 2 AAW36645 Aaw36645 Thrombopo 506
434 56 76.7 19 3 AAB17022 Aab17022 TPO-mimet 507
435 56 76.7 19 4 AAU25864 Aau25864 Human thr 508
436 56 76.7 19 4 AAU25825 Aau25825 Human thr 509
437 56 76.7 19 5 ABB72908 Abb72908 TPO mimet 510
438 56 76.7 19 7 ADJ73060 Adj73060 TPO mimet 511
439 56 76.7 19 8 ADJ52695 Adj52695 CH1 delet 512
440 56 76.7 19 8 ADJ51656 Adj51656 CH1 delet 513
441 56 76.7 22 7 ADN59831 Adn59831 TMP pepti 514
442 56 76.7 22 7 ADN59827 Adn59827 TMP pepti 515
443 56 76.7 25 7 ADN59710 Adn59710 Thrombopo 516
444 56 76.7 25 7 ADN59702 Adn59702 Thrombopo 517
445 56 76.7 30 9 ADY64336 Ady64336 Thrombopo 518
446 56 76.7 42 7 ADN59751 Adn59751 Peptide-v 519
447 54 74.0 18 7 ADN59815 Adn59815 Thrombopo 520
448 54 74.0 18 7 ADN59668 Adn59668 Thrombopo 521
449 54 74.0 18 7 ADN59666 Adn59666 Thrombopo 522
450 54 74.0 22 7 ADN59835 Adn59835 TMP pepti 523
451 54 74.0 22 7 ADN59833 Adn59833 TMP pepti 524
452 54 74.0 25 7 ADN59718 Adn59718 Thrombopo 525
453 54 74.0 25 7 ADN59714 Adn59714 Thrombopo 526
454 53 72.6 15 2 AAW66724 Aaw66724 Peptide c 527
455 53 72.6 15 4 AAU26016 Aau26016 Human thr 528
456 53 72.6 15 8 ADM72480 Adm72480 TPO mimet 529
457 53 72.6 15 8 ADM72481 Adm72481 TPO mimet 530
458 53 72.6 16 4 AAU26017 Aau26017 Human thr 531
459 53 72.6 18 7 ADN59681 Adn59681 Thrombopo 532
460 53 72.6 18 7 ADN59675 Adn59675 Thrombopo 533
461 53 72.6 18 7 ADN59680 Adn59680 Thrombopo 534
462 53 72.6 19 2 AAW09457 Aaw09457 Thrombopo 535

53 72.6 19 2 AAW09492 Aaw09492 Thrombopo 463
53 72.6 19 2 AAW36651 Aaw36651 Thrombopo 464
53 72.6 19 2 AAW33024 Aaw33024 Thrombopo 465
53 72.6 19 2 AAW36643 Aaw36643 Thrombopo 466
53 72.6 19 3 AAB17021 Aab17021 TPO-mimet 467
53 72.6 19 4 AAU25862 Aau25862 Human thr 468
53 72.6 19 4 AAU25870 Aau25870 Human thr 469
53 72.6 19 4 AAU25821 Aau25821 Human thr 470
53 72.6 19 5 ABB72907 Abb72907 TPO mimet 471
53 72.6 19 7 ADJ73059 Adj73059 TPO mimet 472
53 72.6 19 8 ADJ52694 Adj52694 CH1 delet 473
53 72.6 19 8 ADJ51655 Adj51655 CH1 delet 474
53 72.6 25 7 ADN59742 Adn59742 Thrombopo 475
53 72.6 25 7 ADN59730 Adn59730 Thrombopo 476
53 72.6 25 7 ADN59740 Adn59740 Thrombopo 477
53 72.6 43 7 ADN59752 Adn59752 Peptide-v 478
52 71.2 12 3 AAB17309 Aab17309 Synthetic 479
52 71.2 18 7 ADN59653 Adn59653 Thrombopo 480
52 71.2 18 7 ADN59670 Adn59670 Thrombopo 481
52 71.2 18 7 ADN59667 Adn59667 Thrombopo 482
52 71.2 22 7 ADN59820 Adn59820 TMP pepti 483
52 71.2 22 7 ADN59834 Adn59834 TMP pepti 484
52 71.2 22 7 ADN59837 Adn59837 TMP pepti 485
52 71.2 25 7 ADN59689 Adn59689 Thrombopo 486
52 71.2 25 7 ADN59722 Adn59722 Thrombopo 487
52 71.2 25 7 ADN59716 Adn59716 Thrombopo 488
52 71.2 42 7 ADN59818 Adn59818 Peptide- 489
51 69.9 14 8 ADM72511 Adm72511 TPO mimet 490
51 69.9 14 8 ADM72519 Adm72519 TPO mimet 491
51 69.9 15 2 AAW66728 Aaw66728 Peptide c 492
51 69.9 15 4 AAU26033 Aau26033 Human thr 493
51 69.9 15 4 AAU26029 Aau26029 Human thr 494
51 69.9 15 8 ADM72510 Adm72510 TPO mimet 495
51 69.9 15 8 ADM72518 Adm72518 TPO mimet 496
51 69.9 18 7 ADN59654 Adn59654 Thrombopo 497
51 69.9 18 7 ADN59678 Adn59678 Thrombopo 498
51 69.9 18 7 ADN59669 Adn59669 Thrombopo 499
51 69.9 22 7 ADN59836 Adn59836 TMP pepti 500
51 69.9 22 7 ADN59821 Adn59821 TMP pepti 501
51 69.9 25 7 ADN59736 Adn59736 Thrombopo 502
51 69.9 25 7 ADN59691 Adn59691 Thrombopo 503
51 69.9 25 7 ADN59720 Adn59720 Thrombopo 504
50 68.5 10 2 AAW09469 Aaw09469 Thrombopo 505
50 68.5 10 2 AAW36621 Aaw36621 Thrombopo 506
50 68.5 10 2 AAW36772 Aaw36772 Thrombopo 507
50 68.5 10 3 AAB17005 Aab17005 TPO-mimet 508
50 68.5 10 4 AAU25839 Aau25839 Human thr 509
50 68.5 10 4 AAU25840 Aau25840 Human thr 510
50 68.5 10 5 ABB72891 Abb72891 TPO mimet 511
50 68.5 10 7 ADJ73042 Adj73042 TPO mimet 512
50 68.5 10 8 ADJ52677 Adj52677 CH1 delet 513
50 68.5 10 8 ADJ51638 Adj51638 CH1 delet 514
50 68.5 15 2 AAW66725 Aaw66725 Peptide c 515
50 68.5 18 7 ADN59657 Adn59657 Thrombopo 516
50 68.5 22 7 ADN59824 Adn59824 TMP pepti 517
50 68.5 23 7 ADN59775 Adn59775 Peptide-v 518
50 68.5 23 7 ADN59793 Adn59793 Peptide-v 519
50 68.5 25 7 ADN59696 Adn59696 Thrombopo 520
50 68.5 36 7 ADN59763 Adn59763 Peptide-v 521
50 68.5 41 7 ADN59769 Adn59769 Peptide-v 522
50 68.5 46 7 ADN59781 Adn59781 Peptide-v 523
50 68.5 46 7 ADN59787 Adn59787 Peptide-v 524
50 68.5 302 6 ABU33828 Abu33828 Protein e 525
49 67.1 10 2 AAW09472 Aaw09472 Thrombopo 526
49 67.1 10 2 AAW36623 Aaw36623 Thrombopo 527
49 67.1 10 4 AAU25842 Aau25842 Human thr 528
49 67.1 11 2 AAW36791 Aaw36791 Thrombopo 529
49 67.1 13 3 AAB17015 Aab17015 TPO-mimet 530
49 67.1 13 5 ABB72901 Abb72901 TPO mimet 531
49 67.1 13 7 ADJ73054 Adj73054 TPO mimet 532
49 67.1 13 7 ADJ73052 Adj73052 TPO mimet 533
49 67.1 13 7 ADJ73056 Adj73056 TPO mimet 534
49 67.1 13 7 ADJ73053 Adj73053 TPO mimet 535

536	49	67.1	13	7	ADJ73055	Adj73055	TPO mimet	609	47	64.4	18	8	ADQ16693	Adq16693	TPO mimet
537	49	67.1	13	8	ADJ52687	Adj52687	CH1 delet	610	47	64.4	18	9	ADV44431	Adv44431	Anti-teta
538	49	67.1	13	8	ADJ51648	Adj51648	CH1 delet	611	47	64.4	18	9	AEB12903	Aeb12903	Heavy cha
539	49	67.1	14	3	AAB17017	Aab17017	TPO-mimet	612	47	64.4	22	7	ADN59832	Adn59832	TMP pepti
540	49	67.1	14	3	AAB17016	Aab17016	TPO-mimet	613	47	64.4	23	7	ADN59776	Adn59776	Peptide-v
541	49	67.1	14	5	ABB72903	Abb72903	TPO mimet	614	47	64.4	23	7	ADN59794	Adn59794	Peptide-v
542	49	67.1	14	5	ABB72902	Abb72902	TPO mimet	615	47	64.4	25	7	ADN59712	Adn59712	Thrombopo
543	49	67.1	14	8	ADJ52689	Adj52689	CH1 delet	616	47	64.4	25	7	ADN59755	Adn59755	Peptide-v
544	49	67.1	14	8	ADJ52688	Adj52688	CH1 delet	617	47	64.4	36	7	ADN59764	Adn59764	Peptide-v
545	49	67.1	14	8	ADJ51650	Adj51650	CH1 delet	618	47	64.4	41	7	ADN59770	Adn59770	Peptide-v
546	49	67.1	14	8	ADJ51649	Adj51649	CH1 delet	619	47	64.4	46	7	ADN59788	Adn59788	Peptide-v
547	49	67.1	15	3	AAB17018	Aab17018	TPO-mimet	620	47	64.4	46	7	ADN59782	Adn59782	Peptide-v
548	49	67.1	15	5	ABB72904	Abb72904	TPO mimet	621	46	63.0	13	4	AAU26039	Aau26039	Human thr
549	49	67.1	15	8	ADJ52691	Adj52691	CH1 delet	622	46	63.0	13	8	ADM72528	Adm72528	TPO mimet
550	49	67.1	15	8	ADJ52690	Adj52690	CH1 delet	623	46	63.0	18	7	ADN59658	Adn59658	Thrombopo
551	49	67.1	15	8	ADJ51652	Adj51652	CH1 delet	624	46	63.0	18	7	ADN59656	Adn59656	Thrombopo
552	49	67.1	15	8	ADJ51651	Adj51651	CH1 delet	625	46	63.0	22	7	ADN59823	Adn59823	TMP pepti
553	49	67.1	18	7	ADN59672	Adn59672	Thrombopo	626	46	63.0	22	7	ADN59825	Adn59825	TMP pepti
554	49	67.1	18	7	ADN59671	Adn59671	Thrombopo	627	46	63.0	25	4	AAU26042	Aau26042	Human thr
555	49	67.1	18	7	ADN59659	Adn59659	Thrombopo	628	46	63.0	25	7	ADN59694	Adn59694	Thrombopo
556	49	67.1	22	7	ADN59826	Adn59826	TMP pepti	629	46	63.0	25	7	ADN59698	Adn59698	Thrombopo
557	49	67.1	22	7	ADN59838	Adn59838	TMP pepti	630	46	63.0	25	8	ADM72531	Adm72531	TPO mimet
558	49	67.1	22	7	ADN59839	Adn59839	TMP pepti	631	46	63.0	306	5	ABG91624	Abg91624	Purine/py
559	49	67.1	23	7	ADN59795	Adn59795	Peptide-v	632	46	63.0	306	6	ABU36556	Abu36556	Protein e
560	49	67.1	23	7	ADN59777	Adn59777	Peptide-v	633	46	63.0	306	6	ABU34427	Abu34427	Protein e
561	49	67.1	25	7	ADN59724	Adn59724	Thrombopo	634	45	61.6	12	2	AAW35423	Aaw35423	Thrombopo
562	49	67.1	25	7	ADN59700	Adn59700	Thrombopo	635	45	61.6	12	4	AAU26000	Aau26000	Human thr
563	49	67.1	25	7	ADN59744	Adn59744	Thrombopo	636	45	61.6	13	2	AAW09467	Aaw09467	Thrombopo
564	49	67.1	36	7	ADN59765	Adn59765	Peptide-v	637	45	61.6	13	2	AAW35399	Aaw35399	Thrombopo
565	49	67.1	41	7	ADN59771	Adn59771	Peptide-v	638	45	61.6	13	2	AAW35404	Aaw35404	Thrombopo
566	49	67.1	46	7	ADN59789	Adn59789	Peptide-v	639	45	61.6	13	2	AAW35417	Aaw35417	Thrombopo
567	49	67.1	46	7	ADN59757	Adn59757	Peptide-v	640	45	61.6	13	2	AAW33033	Aaw33033	Thrombopo
568	49	67.1	46	7	ADN59783	Adn59783	Peptide-v	641	45	61.6	13	2	AAW35413	Aaw35413	Thrombopo
569	48	65.8	14	8	ADM72513	Adm72513	TPO mimet	642	45	61.6	13	2	AAW35406	Aaw35406	Thrombopo
570	48	65.8	15	4	AAU25834	Aau25834	Human thr	643	45	61.6	13	2	AAW35405	Aaw35405	Thrombopo
571	48	65.8	15	4	AAU26030	Aau26030	Human thr	644	45	61.6	13	2	AAW35422	Aaw35422	Thrombopo
572	48	65.8	15	8	ADM72512	Adm72512	TPO mimet	645	45	61.6	13	2	AAW35397	Aaw35397	Thrombopo
573	48	65.8	18	7	ADN59655	Adn59655	Thrombopo	646	45	61.6	13	4	AAU25994	Aau25994	Human thr
574	48	65.8	18	7	ADN59661	Adn59661	Thrombopo	647	45	61.6	13	4	AAU25997	Aau25997	Human thr
575	48	65.8	18	7	ADN59673	Adn59673	Thrombopo	648	45	61.6	13	4	AAU25991	Aau25991	Human thr
576	48	65.8	22	7	ADN59828	Adn59828	TMP pepti	649	45	61.6	13	4	AAU25990	Aau25990	Human thr
577	48	65.8	22	7	ADN59822	Adn59822	TMP pepti	650	45	61.6	13	4	AAU25984	Aau25984	Human thr
578	48	65.8	22	7	ADN59822	Adn59822	TMP pepti	651	45	61.6	14	2	AAW09466	Aaw09466	Thrombopo
579	48	65.8	23	7	ADN59792	Adn59792	Peptide-v	652	45	61.6	14	2	AAW09462	Aaw09462	Thrombopo
580	48	65.8	23	7	ADN59774	Adn59774	Peptide-v	653	45	61.6	14	2	AAW09465	Aaw09465	Thrombopo
581	48	65.8	23	7	ADN59778	Adn59778	Peptide-v	654	45	61.6	14	2	AAW09482	Aaw09482	Thrombopo
582	48	65.8	23	7	ADN59796	Adn59796	Peptide-v	655	45	61.6	14	2	AAW33031	Aaw33031	Thrombopo
583	48	65.8	25	7	ADN59726	Adn59726	Thrombopo	656	45	61.6	14	2	AAW33031	Aaw33031	Thrombopo
584	48	65.8	25	7	ADN59704	Adn59704	Thrombopo	657	45	61.6	14	2	AAW35407	Aaw35407	Thrombopo
585	48	65.8	25	7	ADN59692	Adn59692	Thrombopo	658	45	61.6	14	2	AAW35408	Aaw35408	Thrombopo
586	48	65.8	36	7	ADN59766	Adn59766	Peptide-v	659	45	61.6	14	2	AAW35398	Aaw35398	Thrombopo
587	48	65.8	36	7	ADN59762	Adn59762	Peptide-v	660	45	61.6	14	2	AAW36633	Aaw36633	Thrombopo
588	48	65.8	36	7	ADN59756	Adn59756	Peptide-v	661	45	61.6	14	2	AAW33029	Aaw33029	Thrombopo
589	48	65.8	40	7	ADN59753	Adn59753	Peptide-v	662	45	61.6	14	2	AAW35396	Aaw35396	Thrombopo
590	48	65.8	41	7	ADN59768	Adn59768	Peptide-v	663	45	61.6	14	2	AAW35401	Aaw35401	Thrombopo
591	48	65.8	41	7	ADN59816	Adn59816	Peptide-	664	45	61.6	14	2	AAW35403	Aaw35403	Thrombopo
592	48	65.8	41	7	ADN59772	Adn59772	Peptide-v	665	45	61.6	14	2	AAW36647	Aaw36647	Thrombopo
593	48	65.8	43	7	ADN59761	Adn59761	Peptide-v	666	45	61.6	14	2	AAW35400	Aaw35400	Thrombopo
594	48	65.8	44	7	ADN59817	Adn59817	Peptide-	667	45	61.6	14	2	AAW35402	Aaw35402	Thrombopo
595	48	65.8	46	7	ADN59780	Adn59780	Peptide-v	668	45	61.6	14	2	AAW33032	Aaw33032	Thrombopo
596	48	65.8	46	7	ADN59790	Adn59790	Peptide-v	669	45	61.6	14	3	AAU25987	Aau25987	TPO-mimet
597	48	65.8	46	7	ADN59786	Adn59786	Peptide-v	670	45	61.6	14	4	AAU25987	Aau25987	Human thr
598	48	65.8	46	7	ADN59784	Adn59784	Peptide-v	671	45	61.6	14	4	AAU25826	Aau25826	Human thr
599	48	65.8	607	3	AAAY58580	Aay58580	Sorangium	672	45	61.6	14	4	AAU25993	Aau25993	Human thr
600	47	64.4	8	2	AAW33037	Aaw33037	Thrombopo	673	45	61.6	14	4	AAU25852	Aau25852	Human thr
601	47	64.4	8	4	AAU25982	Aau25982	Human thr	674	45	61.6	14	4	AAU25989	Aau25989	Human thr
602	47	64.4	8	5	ABP51678	Abp51678	TPO mimet	675	45	61.6	14	4	AAU25983	Aau25983	Human thr
603	47	64.4	8	8	ADQ16692	Adq16692	TPO mimet	676	45	61.6	14	4	AAU25985	Aau25985	Human thr
604	47	64.4	8	9	ADV44430	Adv44430	Agonist.T	677	45	61.6	14	4	AAU25995	Aau25995	Human thr
605	47	64.4	8	9	AEB12902	Aeb12902	TPO mimet	678	45	61.6	14	4	AAU25992	Aau25992	Human thr
606	47	64.4	15	2	AAW66726	Aaw66726	Peptide c	679	45	61.6	14	4	AAU25986	Aau25986	Human thr
607	47	64.4	18	5	ABP51679	Abp51679	TPO mimet	680	45	61.6	14	4	AAU25866	Aau25866	Human thr
608	47	64.4	18	7	ADN59665	Adn59665	Thrombopo	681	45	61.6	14	4	AAU25988	Aau25988	Human thr

682	45	61.6	14	5	ABB72900	Abb72900	TPO mimet	755	42	57.5	15	8	ADM72516	Adm72516	TPO mimet
683	45	61.6	14	7	ADJ73051	Adj73051	TPO mimet	756	42	57.5	18	2	AAW09592	Aaw09592	Thrombopo
684	45	61.6	14	8	ADJ52686	Adj52686	CH1 delet	757	42	57.5	18	2	AAW36743	Aaw36743	Thrombopo
685	45	61.6	14	8	ADJ51647	Adj51647	CH1 delet	758	42	57.5	18	4	AAU25962	Aau25962	Human thr
686	45	61.6	14	8	ADM72515	Adm72515	TPO mimet	759	42	57.5	100	5	ABP47801	Abp47801	Sooty man
687	45	61.6	14	8	ADM72521	Adm72521	TPO mimet	760	42	57.5	110	4	AAM85655	Aam85655	Human imm
688	45	61.6	15	2	AAW66729	Aaw66729	Peptide c	761	42	57.5	126	7	ADM03803	Adm03803	Human pro
689	45	61.6	15	4	AAU26031	Aau26031	Human thr	762	42	57.5	429	5	ABP47758	Abp47758	Protein #
690	45	61.6	15	4	AAU26034	Aau26034	Human thr	763	42	57.5	450	3	AAG50584	Aag50584	Arabidops
691	45	61.6	15	8	ADM72520	Adm72520	TPO mimet	764	42	57.5	475	3	AAG50583	Aag50583	Arabidops
692	45	61.6	15	8	ADM72514	Adm72514	TPO mimet	765	42	57.5	496	9	ABM91663	Abm91663	M. xanthu
693	45	61.6	18	2	AAW09456	Aaw09456	Thrombopo	766	42	57.5	499	3	AAG50582	Aag50582	Arabidops
694	45	61.6	18	2	AAW09487	Aaw09487	Thrombopo	767	42	57.5	512	4	ABG25367	Abg25367	Novel hum
695	45	61.6	18	2	AAW33023	Aaw33023	Thrombopo	768	42	57.5	519	7	ADM05082	Adm05082	Human pro
696	45	61.6	18	2	AAW36638	Aaw36638	Thrombopo	769	41	56.2	18	2	AAW09490	Aaw09490	Thrombopo
697	45	61.6	18	3	AAB17020	Aab17020	TPO-mimet	770	41	56.2	18	2	AAW36641	Aaw36641	Thrombopo
698	45	61.6	18	4	AAU25857	Aau25857	Human thr	771	41	56.2	18	4	AAU25860	Aau25860	Human thr
699	45	61.6	18	4	AAU25820	Aau25820	Human thr	772	41	56.2	18	7	ADN59676	Adn59676	Thrombopo
700	45	61.6	18	5	ABB72906	Abb72906	TPO mimet	773	41	56.2	18	7	ADN59662	Adn59662	Thrombopo
701	45	61.6	18	7	ADJ73058	Adj73058	TPO mimet	774	41	56.2	22	7	ADN59829	Adn59829	TMP pepti
702	45	61.6	18	8	ADJ52693	Adj52693	CH1 delet	775	41	56.2	25	7	ADN59706	Adn59706	Thrombopo
703	45	61.6	18	8	ADJ51654	Adj51654	CH1 delet	776	41	56.2	25	7	ADN59732	Adn59732	Thrombopo
704	45	61.6	19	2	AAW09458	Aaw09458	Thrombopo	777	41	56.2	49	9	ABM94741	Abm94741	M. xanthu
705	45	61.6	19	2	AAW33025	Aaw33025	Thrombopo	778	41	56.2	137	8	ABO84792	Abo84792	Murine ca
706	45	61.6	19	4	AAU25822	Aau25822	Human thr	779	41	56.2	155	3	AAG12015	Aag12015	Arabidops
707	44.5	61.0	316	9	ABM933583	Abm933583	M. xanthu	780	41	56.2	211	3	AAB43611	Aab43611	Human can
708	44	60.3	12	8	ADM72530	Adm72530	TPO mimet	781	41	56.2	211	4	ABG22919	Abg22919	Novel hum
709	44	60.3	13	4	AAU26041	Aau26041	Human thr	782	41	56.2	220	4	ABG18864	Abg18864	Novel hum
710	44	60.3	13	8	ADM72529	Adm72529	TPO mimet	783	41	56.2	291	3	AAG25112	Aag25112	Arabidops
711	44	60.3	14	2	AAW66732	Aaw66732	Peptide c	784	41	56.2	294	3	AAG25111	Aag25111	Arabidops
712	44	60.3	14	4	AAU26040	Aau26040	Human thr	785	41	56.2	294	3	AAG55002	Aag55002	Arabidops
713	44	60.3	15	2	AAW66727	Aaw66727	Peptide c	786	41	56.2	294	3	AAG17087	Aag17087	Arabidops
714	44	60.3	18	2	AAW09497	Aaw09497	Thrombopo	787	41	56.2	306	5	ABG91631	Abg91631	Purine/py
715	44	60.3	18	2	AAW36653	Aaw36653	Thrombopo	788	41	56.2	306	6	ABU35738	Abu35738	Protein e
716	44	60.3	18	2	AAW36648	Aaw36648	Thrombopo	789	41	56.2	308	4	AAG91969	Aag91969	C glutami
717	44	60.3	18	3	AAB17025	Aab17025	TPO-mimet	790	41	56.2	311	7	ABM89301	Abm89301	Rice abio
718	44	60.3	18	4	AAU25872	Aau25872	Human thr	791	41	56.2	312	8	ADT56484	Adt56484	Plant pol
719	44	60.3	18	4	AAU25867	Aau25867	Human thr	792	41	56.2	318	8	ADY22882	Ady22882	Plant ful
720	44	60.3	18	5	ABB72911	Abb72911	TPO mimet	793	41	56.2	326	1	AAP82999	Aap82999	Tobacco G
721	44	60.3	18	7	ADJ73063	Adj73063	TPO mimet	794	41	56.2	338	3	AAG17086	Aag17086	Arabidops
722	44	60.3	18	7	ADN59674	Adn59674	Thrombopo	795	41	56.2	338	3	AAG25110	Aag25110	Arabidops
723	44	60.3	18	8	ADJ52698	Adj52698	CH1 delet	796	41	56.2	364	3	AAG17085	Aag17085	Arabidops
724	44	60.3	18	8	ADJ51659	Adj51659	CH1 delet	797	41	56.2	876	9	ABM92266	Abm92266	M. xanthu
725	44	60.3	22	7	ADN59841	Adn59841	TMP pepti	798	40	54.8	14	2	AAW09484	Aaw09484	Thrombopo
726	44	60.3	23	7	ADN59797	Adn59797	Peptide-v	799	40	54.8	14	2	AAW36766	Aaw36766	Thrombopo
727	44	60.3	23	7	ADN59779	Adn59779	Peptide-v	800	40	54.8	14	2	AAW36635	Aaw36635	Thrombopo
728	44	60.3	25	7	ADN59728	Adn59728	Thrombopo	801	40	54.8	14	3	AAB17009	Aab17009	TPO-mimet
729	44	60.3	35	7	ADN59754	Adn59754	Peptide-v	802	40	54.8	14	4	AAU25854	Aau25854	Human thr
730	44	60.3	36	7	ADN59767	Adn59767	Peptide-v	803	40	54.8	14	5	ABB72895	Abb72895	TPO mimet
731	44	60.3	41	7	ADN59773	Adn59773	Peptide-v	804	40	54.8	14	7	ADJ73046	Adj73046	TPO mimet
732	44	60.3	46	7	ADN59785	Adn59785	Peptide-v	805	40	54.8	14	8	ADJ52681	Adj52681	CH1 delet
733	44	60.3	46	7	ADN59791	Adn59791	Peptide-v	806	40	54.8	14	8	ADJ51642	Adj51642	CH1 delet
734	44	60.3	75	7	ADN59758	Adn59758	Peptide-v	807	40	54.8	19	2	AAW09495	Aaw09495	Thrombopo
735	44	60.3	87	4	AAU63256	Aau63256	Propionib	808	40	54.8	19	2	AAW36646	Aaw36646	Thrombopo
736	44	60.3	87	6	ABM59775	Abm59775	Propionib	809	40	54.8	19	3	AAB17023	Aab17023	TPO-mimet
737	44	60.3	135	4	ABG29134	Abg29134	Novel hum	810	40	54.8	19	4	AAU25865	Aau25865	Human thr
738	44	60.3	200	8	ADN23470	Adn23470	Bacterial	811	40	54.8	19	5	ABB72909	Abb72909	TPO mimet
739	44	60.3	536	6	ABU15399	Abu15399	Protein e	812	40	54.8	19	7	ADJ73061	Adj73061	TPO mimet
740	44	60.3	536	8	ADK13781	Adk13781	E. coli i	813	40	54.8	19	8	ADJ52696	Adj52696	CH1 delet
741	44	60.3	536	8	ADN18093	Adn18093	Bacterial	814	40	54.8	19	8	ADJ51657	Adj51657	CH1 delet
742	43	58.9	11	2	AAW35425	Aaw35425	Thrombopo	815	40	54.8	28	7	AAE38642	Aae38642	Human col
743	43	58.9	11	4	AAU26001	Aau26001	Human thr	816	40	54.8	153	9	ABM92349	Abm92349	M. xanthu
744	43	58.9	18	2	AAW09488	Aaw09488	Thrombopo	817	40	54.8	207	6	ABU34877	Abu34877	Protein e
745	43	58.9	18	2	AAW36639	Aaw36639	Thrombopo	818	40	54.8	239	8	ADM80079	Adm80079	Spiramyci
746	43	58.9	18	4	AAU25858	Aau25858	Human thr	819	40	54.8	239	8	ADN97595	Adn97595	S ambofac
747	43	58.9	18	7	ADN59679	Adn59679	Thrombopo	820	40	54.8	261	7	ADH86920	Adh86920	Enterococ
748	43	58.9	18	7	ADN59677	Adn59677	Thrombopo	821	40	54.8	271	8	ADS42097	Ads42097	Bacterial
749	43	58.9	25	7	ADN59734	Adn59734	Thrombopo	822	40	54.8	311	4	AAM41798	Aam41798	Human pol
750	43	58.9	25	7	ADN59738	Adn59738	Thrombopo	823	40	54.8	329	4	ABG23127	Abg23127	Novel hum
751	43	58.9	40	7	ADN59760	Adn59760	Peptide-v	824	40	54.8	355	6	ABR58696	Abr58696	Human can
752	43	58.9	285	6	ABU48553	Abu48553	Protein e	825	40	54.8	355	7	ADF13714	Adf13714	Tumor-Ass
753	42	57.5	14	8	ADM72517	Adm72517	TPO mimet	826	40	54.8	355	8	ADS00100	Ads00100	Human ocu
754	42	57.5	15	4	AAU26032	Aau26032	Human thr	827	40	54.8	355	8	ADS00098	Ads00098	Human ocu

828	40	54.8	355	9	ADW14773	Adw14773 Tumor-ass	901	38	52.1	2	AAW36624	Aaw36624 Thrombopo
829	40	54.8	410	6	ABU22259	Abu22259 Protein e	902	38	52.1	4	AAU25843	Aau25843 Human thr
830	40	54.8	410	6	ABU20206	Abu20206 Protein e	903	38	52.1	89	AAM83690	Aam83690 Human imm
831	40	54.8	486	6	ABP77210	Abp77210 N. gonorr	904	38	52.1	126	AAy76522	Aay76522 Human ova
832	40	54.8	492	7	ABO82111	Abo82111 Pseudomon	905	38	52.1	132	ADM04387	Adm04387 Human pro
833	40	54.8	513	2	AAy38789	Aay38789 Neisseria	906	38	52.1	161	AAG28444	Aag28444 Arabidops
834	40	54.8	513	2	AAy38786	Aay38786 Neisseria	907	38	52.1	162	ADx94945	Adx94945 Plant ful
835	40	54.8	513	2	AAy38787	Aay38787 Neisseria	908	38	52.1	174	ADx91393	Adx91393 Plant ful
836	40	54.8	513	2	AAy38788	Aay38788 Neisseria	909	38	52.1	206	AAG28443	Aag28443 Arabidops
837	40	54.8	513	9	AEb49392	Aeb49392 N. mening	910	38	52.1	212	ABO68391	Abo68391 Pseudomon
838	40	54.8	513	9	AEb49394	Aeb49394 N. mening	911	38	52.1	232	ADx95040	Adx95040 Plant ful
839	40	54.8	513	9	AEb49395	Aeb49395 N. gonorr	912	38	52.1	240	ADy10154	Ady10154 Plant ful
840	40	54.8	513	9	AEb49397	Aeb49397 N. gonorr	913	38	52.1	243	ADJ71135	Adj71135 Human hea
841	40	54.8	514	7	ABO82206	Abo82206 Pseudomon	914	38	52.1	244	ABU23180	Abu23180 Protein e
842	40	54.8	549	2	AAW40113	Aaw40113 Human alp	915	38	52.1	270	ADx92531	Adx92531 Plant ful
843	40	54.8	603	8	ADS25043	Ads25043 Bacterial	916	38	52.1	310	ABM96497	Abm96497 M. xanthu
844	40	54.8	656	8	ADJ71949	Adj71949 Human PMM	917	38	52.1	312	ABM67747	Abm67747 Photorhab
845	40	54.8	686	8	ABM84061	Abm84061 Human dia	918	38	52.1	313	ADf04965	Adf04965 Bacterial
846	40	54.8	699	9	ADv85513	Adv85513 Human tra	919	38	52.1	315	ABU21946	Abu21946 Protein e
847	40	54.8	721	9	ADZ26477	Adz26477 Human ery	920	38	52.1	321	ADx65940	Adx65940 Plant ful
848	40	54.8	994	8	ABM82875	Abm82875 Human dia	921	38	52.1	327	ADx91347	Adx91347 Plant ful
849	40	54.8	1061	8	ABM82874	Abm82874 Human dia	922	38	52.1	333	ADx68901	Adx68901 Plant ful
850	40	54.8	1078	4	AAm40012	Aam40012 Human pol	923	38	52.1	334	ABO65391	Abo65391 Klebsiell
851	40	54.8	1712	5	AAO17361	Aao17361 Human alp	924	38	52.1	334	ADx95928	Adx95928 Plant ful
852	40	54.8	1712	8	ADS10473	Ads10473 Human the	925	38	52.1	338	ADt57351	Adt57351 Plant pol
853	40	54.8	1712	8	ADU06526	Adu06526 Novel bro	926	38	52.1	338	ADy09292	Ady09292 Plant ful
854	39.5	54.8	1712	9	ADY59885	Ady59885 Human col	927	38	52.1	339	ADs14943	Ads14943 Pseudomon
855	39.5	54.1	821	4	ABG29052	Abg29052 Novel hum	928	38	52.1	343	ADx96259	Adx96259 Plant ful
856	39.5	54.1	821	4	ABG29013	Abg29013 Novel hum	929	38	52.1	344	ADY10815	Ady10815 Plant ful
857	39.5	54.1	821	4	ABG29872	Abg29872 Novel hum	930	38	52.1	344	ADx95200	Adx95200 Plant ful
858	39	53.4	14	3	ABB17010	Abb17010 TPO-mimet	931	38	52.1	344	ADx96519	Adx96519 Plant ful
859	39	53.4	14	5	ABB72896	Abb72896 TPO mimet	932	38	52.1	344	ADY08650	Ady08650 Plant ful
860	39	53.4	14	7	ADJ73047	Adj73047 TPO mimet	933	38	52.1	344	ADY13023	Ady13023 Plant ful
861	39	53.4	14	8	ADJ52682	Adj52682 CH1 delet	934	38	52.1	344	ADY09844	Ady09844 Plant ful
862	39	53.4	14	8	ADJ51643	Adj51643 CH1 delet	935	38	52.1	344	ADY12363	Ady12363 Plant ful
863	39	53.4	18	2	AAW09489	Aaw09489 Thrombopo	936	38	52.1	344	ADY25241	Ady25241 Plant ful
864	39	53.4	18	2	AAW36640	Aaw36640 Thrombopo	937	38	52.1	344	ADx89992	Adx89992 Plant ful
865	39	53.4	18	4	AAU25859	Aau25859 Human thr	938	38	52.1	344	ADY11386	Ady11386 Plant ful
866	39	53.4	95	5	ABP06014	Abp06014 Human ORF	939	38	52.1	345	ADY12305	Ady12305 Plant ful
867	39	53.4	95	5	AAU09166	Aau09166 Human cyt	940	38	52.1	349	ADx76194	Adx76194 Plant ful
868	39	53.4	100	5	ABP47800	Abp47800 HIV-2prot	941	38	52.1	350	AAW60244	Aaw60244 Amino aci
869	39	53.4	128	4	AAU57127	Aau57127 Propionib	942	38	52.1	350	AAU69755	Aau69755 Escherich
870	39	53.4	128	6	ABM53646	Abm53646 Propionib	943	38	52.1	351	ADx95138	Adx95138 Plant ful
871	39	53.4	168	4	ABB65448	Abb65448 Drosophil	944	38	52.1	351	ADY06579	Ady06579 Plant ful
872	39	53.4	201	8	ADN25936	Adn25936 Bacterial	945	38	52.1	355	ADY06836	Ady06836 Plant ful
873	39	53.4	230	2	AAR60767	Aar60767 HIV-2 ROD	946	38	52.1	355	ADx93406	Adx93406 Plant ful
874	39	53.4	267	5	AAU91162	Aau91162 Bordella	947	38	52.1	356	ADY08052	Ady08052 Plant ful
875	39	53.4	318	9	ABM91698	Abm91698 M. xanthu	948	38	52.1	356	ADx66849	Adx66849 Plant ful
876	39	53.4	321	8	ADx65923	Adx65923 Plant ful	949	38	52.1	357	ADx67086	Adx67086 Plant ful
877	39	53.4	324	8	ADY22836	Ady22836 Bacterial	950	38	52.1	357	ADx68311	Adx68311 Plant ful
878	39	53.4	397	6	ABR82135	Abr82135 Bacterial	951	38	52.1	359	ADx68311	Adx68311 Plant ful
879	39	53.4	412	7	ABM87926	Abm87926 Rice abio	952	38	52.1	361	ADx67626	Adx67626 Plant ful
880	39	53.4	428	5	ABP47757	Abp47757 Protein #	953	38	52.1	362	ADx94971	Adx94971 Plant ful
881	39	53.4	476	8	ADT58335	Adt58335 Plant pol	954	38	52.1	362	ADx93509	Adx93509 Plant ful
882	39	53.4	497	8	ADT57737	Adt57737 Plant pol	955	38	52.1	362	ADY11959	Ady11959 Plant ful
883	39	53.4	517	8	ADY23040	Ady23040 Plant ful	956	38	52.1	363	ADx93520	Adx93520 Plant ful
884	39	53.4	526	8	ADY07792	Ady07792 Plant ful	957	38	52.1	364	ADx93884	Adx93884 Plant ful
885	39	53.4	619	7	ADH87640	Adh87640 Enterococ	958	38	52.1	364	ADY13598	Ady13598 Plant ful
886	39	53.4	706	6	ABU22341	Abu22341 Protein e	959	38	52.1	364	ADY06580	Ady06580 Plant ful
887	39	53.4	764	2	AAW81784	Aaw81784 P. putida	960	38	52.1	364	ADY22675	Ady22675 Plant ful
888	39	53.4	1014	1	AAP80810	Aap80810 Sequence	961	38	52.1	372	AAg80999	Aag80999 HIV prote
889	39	53.4	1027	1	AAP81773	Aap81773 Sequence	962	38	52.1	373	ADY10783	Ady10783 Plant ful
890	39	53.4	1035	2	AAR04025	Aar04025 Pol gene	963	38	52.1	384	ADA36197	Ada36197 Acinetoba
891	39	53.4	1036	2	AAR20599	Aar20599 ROD HIV-2	964	38	52.1	489	ABO76974	Abo76974 Pseudomon
892	39	53.4	1036	3	AAy51978	Aay51978 HIV-2 ROD	965	38	52.1	496	ABG99875	Abg99875 S. cinnam
893	39	53.4	1036	6	ADA07961	Ada07961 HIV-2 pol	966	38	52.1	572	ABO80566	Abo80566 Pseudomon
894	39	53.4	1036	8	ADL23699	Adl23699 HIV-2 ROD	967	38	52.1	694	ADQ67654	Adq67654 Novel hum
895	39	53.4	1256	7	ADE97412	Ade97412 Murine ne	968	38	52.1	694	ABb57877	Abb57877 Drosophil
896	39	53.4	1275	6	ABJ25467	Abj25467 Aspergill	969	38	52.1	904	ABU34516	Abu34516 Protein e
897	39	53.4	1435	6	ABJ26067	Abj26067 Aspergill	970	38	52.1	904	ABU36656	Abu36656 Protein e
898	39	53.4	2773	8	ADN27223	Adn27223 Bacterial	971	38	52.1	905	AAg91240	Aag91240 C glutami
899	39	53.4	3437	4	ABB66639	Abb66639 Drosophil	972	38	52.1	905	ADD13196	Add13196 C. glutam
900	38	52.1	10	2	AAW09473	Aaw09473 Thrombopo	973	38	52.1	999	ADx40028	Adx40028 HIV Pol p

974 Ade97411 Norway ra
975 Aab84865 Murine ne
976 Abb71980 Drosophil
977 Abo67862 Pseudomon
978 Abo76480 Pseudomon
979 Abo74914 Pseudomon
980 Aaw35427 Thrombopo
981 Aau26002 Human thr
982 Aaw09479 Thrombopo
983 Aaw09481 Thrombopo
984 Aaw36630 Thrombopo
985 Aaw36632 Thrombopo
986 Aaw36631 Thrombopo
987 Aab17013 TPO-mimet
988 Aau25849 Human thr
989 Aau25850 Human thr
990 Aau25851 Human thr
991 Abb72899 TPO mimet
992 Adj73050 TPO mimet
993 Adj52685 CH1 delet
994 Adj51646 CH1 delet
995 Aau54078 Propionib
996 Abm50597 Propionib
997 Abp47799 SIV prote
998 Aam84222 Human imm
999 Aau57517 Propionib
1000 Abm54036 Propionib

ALIGNMENTS

RESULT 1
AAW09463
ID AAW09463 standard; protein; 14 AA.
XX AAW09463;
DT 10-SEP-1997 (first entry)
XX Thrombopoietin receptor binding compound peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .14
FT /note= "Preferably linkages are selected from: -
FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
FT lower alkyl"
FT Modified-site 1
FT /note= "Preferably N-terminus is selected from: -NRR1; -
FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
FT benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3
FT substitutions on the phenyl ring selected from lower
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
FT independently selected from hydrogen and lower alkyl"
FT Modified-site 14
FT /note= "Preferably C-terminus is -C(O)R2 where R2 is
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
FT and R4 are independently selected from hydrogen and lower
FT alkyl, and where the nitrogen atom of the -NR3R4 group
FT can optionally be the amine group of the N-terminus of
FT the peptide forming a cyclic peptide"
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX (GLAX) GLAXO GROUP LTD.
PA Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
PI WPI; 1997-051883/05.
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX Claim 18; Page 89; 106pp; English.
PS The present sequence is a compound which binds to thrombopoietin (TPO)
XX receptor (TR). It has a molecular weight of < 8000 Da, and a binding
CC affinity to TR as expressed by an IC50 of no more than about 100 mum. The
CC compound (especially if modified, see features table) can be used for
CC treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAAARA 14
Db |||||
1 IEGPTLRQWLAAARA 14
RESULT 2
AAW09468
ID AAW09468 standard; protein; 14 AA.
XX AAW09468;
AC AAW09468;
XX
DT 10-SEP-1997 (first entry)
XX Thrombopoietin receptor binding compound peptide (part of a dimer).
DE Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.
KW Synthetic.
KW
XX
OS
XX
FH Key Location/Qualifiers
FT Cross-links 14
FT /note= "Linked to the omega Lys from AAW19534"
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.

CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14

RESULT 5
AAW36774
ID AAW36774 standard; peptide; 14 AA.
XX
AC AAW36774;
XX
DT 11-MAR-1998 (first entry)
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 14
-FT /note= "NH2-Ala"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 77; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14

RESULT 6
ADI24843
ID ADI24843 standard; peptide; 14 AA.
XX
AC ADI24843;
XX
DT 15-APR-2004 (first entry)
XX
DE AF 12505 as active moiety for pharmacologically active peptide.
XX
KW pharmacologically active peptide conjugate; enzymatic cleavage; pain;
KW HIV; cancer; diabetes; incontinence; hypertension; amnesia;
KW Alzheimer's disease; fever; depression; sex hormone regulation;
KW eating disorder; schizophrenia; osteoporosis; insomnia;
KW Central nervous system disorder; contraceptive.
XX
OS Synthetic.
XX
PN WO9946283-A1.
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1999; 99WO-DK000118.
XX
PR 09-MAR-1998; 98DK-00000317.
XX
PA (ZEAL-) ZEALAND PHARM AS.
XX
PI Larsen BD;
XX
DR WPI; 1999-561659/47.
XX
PT New peptide conjugates used for treating, e.g. pain, HIV, depression,
PT schizophrenia, osteoporosis or insomnia.
XX
PS Claim 24; Page 90; 113pp; English.
XX
CC The invention relates to a novel pharmacologically active peptide
CC conjugate having a reduced tendency towards enzymatic cleavage comprises
CC X and Z, where: (a) X is a pharmacologically active peptide sequence; and
CC (b) Z is a stabilising peptide sequence of 4-20 amino acid units
CC covalently bound to X, where each amino acid unit in the stabilizing
CC peptide sequence, Z being selected from Ala, Leu, Ser, Thr, Tyr, Asn,
CC Gln, Asp, Glu, Lys, Arg, His, Met, Orn, and amino acid units of formula -
CC NH-C(R1)(R2)-C(=O)- (I), where: R1 and R2 are H, 1-6C alkyl, phenyl, and
CC phenyl-methyl, where 1-6C-alkyl is optionally substituted with 1 -3
CC substituents selected from halogen, hydroxy, amino, cyano, nitro,
CC sulfono, and carboxy, and phenyl and phenyl-methyl are optionally
CC substituted with 1-3 substituents selected from 1-6C-alkyl, 2-6C-alkenyl,
CC halogen, hydroxy, amino, cyano, nitro, sulfono, and carboxy, or R 1 and R
CC 2 together with the C atom to which they are bound form a cyclopentyl,
CC cyclohexyl or cycloheptyl ring, e.g. 2,4-diaminobutanoic acid and 2,3-
CC diaminopropanoic acid; the ratio between the half-life of the peptide
CC conjugate and the half-life of the corresponding active peptide sequence,
CC X, when treated with carboxypeptidase A or leucine aminopeptidase in
CC about 50 mM phosphate buffer solution at about pH 7.4 and about 37 deg C
CC or in serum or plasma is at least about 2 (preferably at least about 10),
CC or when the pharmacologically active peptide X is not orally absorbed,
CC the conjugate is adsorbed, or a salt, with the proviso that the
CC pharmacologically active peptide conjugate is not selected from sequences
CC (ADI24837)-(ADI24841). The peptide conjugates can be used for treating
CC e.g. pain, HIV, cancer, diabetes, incontinence, hypertension, amnesia,
CC Alzheimer's disease, fever, depression, sex hormone regulation, eating
CC disorders, schizophrenia, osteoporosis or insomnia. They can also be used
CC for treating e.g. CNS disorders and as contraceptives. The conjugated
CC peptides are less susceptible to degradation by proteases compared to the
CC corresponding free pharmacologically active peptides. This sequence
CC represents a pharmacologically active peptide as the X part of the

CC peptide of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14
RESULT 7
AAY96515
ID AAY96515 standard; peptide; 14 AA.
XX
AC AAY96515;
XX
DT 04-SEP-2000 (first entry)
XX
XX Thrombopoietin mimetic peptide.
DE Thrombopoietin; mimetic; TMP; TPO; platelet; megakaryocyte; production;
XX Thrombopoietin; mimetic; TMP; TPO; platelet; megakaryocyte; production;
KW anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;
KW immunosuppressive; anti-inflammatory.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 14
FT /note= "subunits in the dimer are covalently bonded at
FT each carboxy terminus through peptide linkage with NH2-
FT (CH2)4-CH(CONH2)-NH-CO-(CH2)2-NH2"
XX
PN WO200024770-A2.
XX
PD 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US024834.
XX
PR 23-OCT-1998; 98US-0105348P.
XX (AMGE-) AMGEN INC.
XX
PI Liu C, Feige U, Cheetham J;
XX WPI; 2000-365108/31.
DR
XX Thrombopoietic peptides which activate mpl receptors and increase the
PT production of platelets or platelet precursors, useful for treatment of
PT diseases which involve thrombocytopenia.
XX
PS Claim 7; Page 60; 91pp; English.
XX
CC A compound which binds to an mpl receptor comprising a thrombopoietin
CC mimetic peptide (TMP) dimer joined by a linker [TMP₁-(L₁)_nTMP₂], is
CC new. TMP₁ and TMP₂ are amino acid sequences varying from at least 10 to
CC 14 residues in length comprising X₂-X₁⁰, X₂-X₁¹, X₂-X₁², X₂-
CC X₁³, X₂-X₁⁴, X₁-X₁⁰, X₁-X₁¹, X₁-X₁², X₁-X₁³, and X₁-
CC X₁⁴. X₁¹ = I, A, V, L, S or R; X₂² = E, D, K or V; X₃ = G or A; X₄ =
CC P; X₅ = T or S; X₆ = L, I, V, A or F; X₇ = R or K; X₈ = Q, N, or E;
CC X₉ = W, Y or F; X₁₀ = L, I, V, A, F, M, or K; X₁₁ = A, I, V, L, F,
CC S, T, K, H, or G; X₁₂ = A, I, V, L, F, G, S, or Q; X₁₃ = R, K, T, V,
CC N, Q or G; X₁₄ = A, I, V, L, F, T, R, E, or G; L₁ = linker comprising
CC 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate
CC the c-Mpl receptor which mediates the activity of endogenous
CC thrombopoietin. The TMPs are useful for increasing the production of
CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which
CC is useful for treatment of diseases which involve thrombocytopenia, e.g.
CC aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency
CC virus associated ITP, and systemic lupus erythematosus

SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14
RESULT 8
AAB16962
ID AAB16962 standard; peptide; 14 AA.
XX
AC AAB16962;
XX
DT 31-OCT-2000 (first entry)
XX
XX TPO-mimetic peptide TMP SEQ ID NO:13.
DE
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 189; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 3; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.6e-05; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 9
AAU25827
ID AAU25827 standard; peptide; 14 AA.
XX
AC AAU25827;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #13.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 69-70; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 10
AAU26004
ID AAU26004 standard; peptide; 14 AA.
XX
AC AAU26004;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #190.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 147; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 11

ABB72853
ID ABB72853 standard; peptide; 14 AA.

XX
AC ABB72853;

XX
DT 05-APR-2002 (first entry)

XX
DE TPO mimetic peptide SEQ ID NO:13.

XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX
OS Homo sapiens.

OS Synthetic.

XX
PN WO200183525-A2.

XX
PD 08-NOV-2001.

XX
PF 02-MAY-2001; 2001WO-US014310.

XX
PR 03-MAY-2000; 2000US-00563286.

XX
PA (AMGE-) AMGEN INC.

XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX
DR WPI; 2002-130313/17.

XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX
PS Claim 39; Page 43; 176pp; English.

CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14

Db 1 IEGPTLRQWLAARA 14

RESULT 12

ABP51669
ID ABP51669 standard; peptide; 14 AA.

XX
AC ABP51669;

XX
DT 01-OCT-2002 (first entry)

XX
DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:1.

XX
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

OS Homo sapiens.

OS Synthetic.

XX
PN WO200246238-A2.

XX
PD 13-JUN-2002.

XX
PF 05-DEC-2001; 2001WO-US047656.

XX
PR 05-DEC-2000; 2000US-0251448P.

XX
PR 04-MAY-2001; 2001US-0288889P.

XX
PR 29-MAY-2001; 2001US-0294068P.

XX
PA (ALEX-) ALEXION PHARM INC.

XX
PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX
DR WPI; 2002-566610/60.

XX
PT A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoetin mimetic.

XX
PS Claim 18; Page 6; 113pp; English.

XX
CC The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC promegakaryocytes or megakaryocytes, which results in increased platelet
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
CC the exemplification of the present invention

XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 13
AAE18011
ID AAE18011 standard; peptide; 14 AA.

XX
AC AAE18011;
DT 07-MAY-2002 (first entry)

XX Human ligand #31 attached to an adenoviral vector.

DE Human; adenoviral coat protein; non-native ligand; cell-surface receptor;
KW therapy; anti-tumour agent; tumour necrosis factor; cancer; brain; lung;
KW ovary; breast; prostate.

XX Homo sapiens.

OS

XX WO200192549-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017391.

XX 31-MAY-2000; 2000US-0208451P.

XX 02-AUG-2000; 2000US-00631191.

XX (GENV-) GENVEC INC.

XX Wickham TJ, Kovesdi I, Roelvink PW, Einfeld D, Brough DE;

XX Lizonova A;

XX WPI; 2002-147620/19.

XX Adenoviral coat protein which permits production of adenoviral vectors

XX that bind and infect host cells not naturally infected by adenovirus,

XX comprises various non-native ligands.

XX Claim 4; Page 45; 45pp; English.

XX The invention relates to adenoviral coat proteins comprising various non-
XX native ligands. The invention provides a method of controlled gene
XX expression utilising selectively replication competence and also a method
XX and a composition for targetting an adenoviral vector. A system
XX comprising a cell having a non-native cell-surface receptor, and a virus
XX having a non-native ligand which binds the non-native cell-surface
XX receptor of the cell is useful for propagating a virus and also for
XX assaying gene function. The system is also useful for isolating a nucleic
XX acid encoding a product comprising a desired property. Further the system
XX is useful for identifying functionally related coding sequences.

XX Adenoviral vector comprising a non-native nucleic acid encoding a
XX therapeutic agent such as anti-tumour agent, preferably tumour necrosis
XX factor and a second non-native nucleic acid encoding an agent that
XX facilitates imaging and a targetting agent is useful for treating an
XX animal. The therapeutic agent can be used to treat cancer of the brain,
XX lung, ovary, breast and prostate. The present sequence is human non-
XX native ligand attached to an adenoviral vector

XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 14
ABG71747
ID ABG71747 standard; peptide; 14 AA.

XX
AC ABG71747;

XX 20-JAN-2003 (first entry)

XX TPO receptor, MPL, agonist peptide, AP12505.

DE Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
KW complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
KW thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
KW T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
KW proliferation; growth; differentiation; haematopoietic cell;
KW platelet progenitor cell; immune disorder; thrombocytopenia;
KW disseminated intravascular coagulation; stem cell; transplantation;
KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.

XX Synthetic.

OS WO200278612-A2.

XX 10-OCT-2002.

XX 02-APR-2002; 2002WO-US010301.

XX 02-APR-2001; 2001US-0281183P.

XX (PURD) PURDUE PHARMA LP.

XX Soltis DA, Burch RM, Ogert RA;

XX WPI; 2003-040615/03.

XX New thrombopoietin synthebodies, useful for stimulating proliferation,
XX growth, or differentiation of hematopoietic cells, for treating or
XX preventing hematopoietic or immune disorders, e.g. thrombocytopenia.

XX Claim 62; Page 71; 97pp; English.

XX The invention discloses a variant of an immunoglobulin (Ig) variable
XX heavy or light chain domain that comprises at least one complementarity
XX determining region (CDR) and framework regions flanking the CDR. The CDR
XX also has added or substituted to it, at least one binding sequence which
XX is heterologous to the CDR and is an antigenic, agonistic sequence from a
XX thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
XX can be a binding sequence heterologous to the CDR, a cytotoxic T-
XX lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
XX cell sequence or a combination of each. The variant or thrombopoietin
XX synthebody, pharmaceutical and vaccine compositions are useful for
XX stimulating proliferation, growth or differentiation of haematopoietic
XX cells, particularly platelet progenitor cells. The variants are also
XX useful for treating or preventing haematopoietic or immune disorders
XX resulting from chemotherapy, radiation therapy, or bone marrow
XX transfusions (e.g. thrombocytopenia or disseminated intravascular
XX coagulation). Compositions comprising the synthebodies can be used for
XX the mobilisation, amplification and ex vivo expansion of stem cells and
XX committed precursor cells for autologous and allogeneic transplantation
XX as well as for the expansion of stem cells for gene therapy. They are
XX also useful as diagnostic or analytical reagents for studying the

CC function of thrombopoietin and its receptor in vivo or in vitro. The
CC sequence presented is the TPO receptor (MPL) agonist peptide, AF12505
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 15
ABR62907
ID ABR62907 standard; peptide; 14 AA.
XX
AC ABR62907;
XX
DT 04-DEC-2003 (first entry)
XX
DE Thrombopoietin mimetic peptide AF12505.
XX
KW Thrombopoietin; mimetic; thrombocytopaenia; antibody targeting.
XX
OS Synthetic.
XX
PN WO2003059251-A2.
XX
PD 24-JUL-2003.
XX
PF 22-OCT-2002; 2002WO-US033991.
XX
PR 22-OCT-2001; 2001US-0344614P.
PR 19-SEP-2002; 2002US-0412455P.
XX
XX (SCRI) SCRIPPS RES INST.
XX
XX Barbas CF, Rader C, Sinha SC, Lerner R;
PI
XX
XX WPI; 2003-636673/60.
DR
XX
XX Antibody targeting compound useful e.g. for diagnostic immunoassays and
PT treating microbial diseases comprises targeting or biological agent
PT covalently linked to combining site of the antibody.
XX
PS Example 7; Page 62; 56pp; English.
XX
XX The present sequence is that of thrombopoietin (TPO) mimetic peptide
CC AF12505, which mimics the activity of recombinant TPO. The invention
CC provides antibody targeting compounds that are used to reprogram the
CC specificity of an antibody. The antibody targeting compound is linked to
CC the combining site of the antibody, such that the modified antibody takes
CC on the binding specificity of the targeting agent. In an example from the
CC invention, a TPO receptor targeting antibody compound was prepared by
CC covalently linking peptide ABL2505 to aldolase monoclonal antibody 38C2.
CC The TPO receptor targeting antibody compound can be used to treat
CC thrombocytopaenia resulting from chemotherapy and bone marrow
CC transplantation
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 16

ADC33697
ID ADC33697 standard; peptide; 14 AA.
XX
AC ADC33697;
XX
DT 18-DEC-2003 (first entry)
XX
DE Erythropoietin receptor/erythropoietin consensus peptide SEQ ID NO:1.
XX
KW chimeric retrovirus envelope protein; ecotropic envelope protein;
KW cytotstatic; gene therapy; cancer.
XX
OS Synthetic.
XX
PN WO2003076596-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US007323.
XX
PR 08-MAR-2002; 2002US-0362655P.
XX
PA (UYMA-) UNIV MASSACHUSETTS.
XX
XX Green MR, Gollan TJ;
PI
XX WPI; 2003-722332/68.
DR
XX
XX New chimeric retrovirus envelope protein comprising an ecotropic envelope
PT protein and a heterologous short peptide ligand inserted within the
PT ecotropic envelope protein useful for treating cancer.
XX
PS Disclosure; SEQ ID NO 1; 42pp; English.
XX
XX The present invention describes a chimeric retrovirus envelope protein
CC (I) comprising an ecotropic envelope protein and a heterologous short
CC peptide ligand inserted within the ecotropic envelope protein. Also
CC described: (1) a nucleic acid molecule comprising a sequence encoding the
CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic
CC acid sequence encoding the chimeric envelope protein; (3) a recombinant
CC retroviral particle comprising a chimeric envelope protein comprising a
CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)
CC identifying a nucleic acid sequence encoding the chimeric envelope
CC protein that alters viral tropism; (5) delivering a nucleic acid sequence
CC to a cell; and (6) treating cancer. (I) has cytotstatic activity and can
CC be used in gene therapy. The chimeric retrovirus envelope protein is
CC useful for treating cancer, which comprises providing a cancer cell, e.g.
CC human cancer cell and infecting the cancer cell with a virus, e.g.
CC retrovirus comprising the chimeric envelope protein comprising a
CC heterologous short peptide ligand and a therapeutically useful gene, e.g.
CC encoding thymidine kinase. The present sequence represents an
CC erythropoietin receptor/erythropoietin consensus peptide, which is given
CC in the exemplification of the present invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 17
ADN59652
ID ADN59652 standard; peptide; 14 AA.
XX
AC ADN59652;
XX
DT 01-JUL-2004 (first entry)
XX

DE Thrombopoietin mimetic peptide (TMP), seq id 1.
XX
KW Haemostatic; antianaemic; immunosuppressive; platelet;
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;
KW autoimmune haemolytic anaemia; Hughe's syndrome;
KW lupoid thrombocytopaenia.
XX
OS Homo sapiens.
XX
XX WO2003031589-A2.
PN
XX
PD 17-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032552.
PF
XX
XX 11-OCT-2001; 2001US-0328666P.
PR
PR 10-OCT-2002; 2002US-00269806.
XX
XX
PA (AMGE-) AMGEN INC.
XX
XX Min H, Sitney KC, Hartley C;
PI
XX WPI; 2003-403101/38.
DR
XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
PT which stimulate the production of platelets and/or the production of
PT platelet precursors, useful for treating thrombocytopenia.
XX
PS Disclosure; SEQ ID NO 1; 126pp; English.
XX
CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of
CC platelets and/or the production of platelet precursors, is new. Further
CC disclosed is a composition of matter (II) that binds to an mpl receptor,
CC and a pharmaceutical composition comprising (II) and a carrier. The
CC pharmaceutical composition of the invention is useful for treating
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or
CC platelets in a patient. The TMP of the invention is useful for treating
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,
CC autoimmune haemolytic anaemia, Hughe's syndrome and lupoid
CC thrombocytopaenia. The TMP of the invention is also useful for
CC maintaining the viability or storage life of platelets and/or
CC megakaryocytes and its derived cells. The compounds demonstrate an
CC improved ability to bind to and/or trigger transmembrane signal through,
CC i.e. activating, the mpl receptor the compounds have superior
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in
CC vitro, the production of platelets and/or megakaryocytopenic activity,
CC i.e. the ability to stimulate, in vivo and in vitro, the production of
CC platelet precursors. Further, certain of the compounds also exhibit
CC superior therapeutic properties, such as improved plasma half-life,
CC biological activity and in vivo circulation time. The current sequence
CC represents a TMP of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
RESULT 18
ADL27293
ID ADL27293 standard; protein; 14 AA.
XX
AC ADL27293;
Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
RESULT 19
ADM72483
ID ADM72483 standard; peptide; 14 AA.
XX
AC ADM72483;
XX
XX 17-JUN-2004 (first entry)
DT
XX TPO mimetic peptide fragment.
DE
XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
XX WO2004026332-A1.
PN
XX

XX 03-JUN-2004 (first entry)
DT
XX Amino acid sequence of a thrombopoietin agonist peptide.
DE
XX fusion protein; C4bp; alpha chain; systemic lupus erythematosus.
KW
XX Homo sapiens.
OS
OS Synthetic.
XX
XX WO2004020639-A2.
PN
XX 11-MAR-2004.
PD
XX 12-AUG-2003; 2003WO-EP008928.
PF
XX 14-AUG-2002; 2002EP-00292043.
PR
XX (AVID-) AVIDIS SA.
PA
XX Garnier L, Hill F, Julien M;
PI
XX WPI; 2004-239202/22.
DR
XX Obtaining a recombinant fusion protein, useful for treating lupus,
PT comprises providing a prokaryotic host cell carrying a nucleic acid
PT encoding the recombinant protein operably linked to a promoter functional
PT in the prokaryotic cell.
XX
PS Claim 8; Page 48; 69pp; English.
XX
CC The specification describes a method for obtaining a recombinant fusion
CC protein comprising a scaffold of a C-terminal core protein of C4bp alpha
CC chain, where the recombinant fusion protein is capable of forming
CC multimers in soluble form in a prokaryotic host cell. The method
CC comprises providing a prokaryotic host cell carrying a nucleic acid
CC encoding the recombinant protein operably linked to a promoter functional
CC in the prokaryotic cell, culturing the host cell under conditions where
CC the recombinant protein is expressed, and recovering the recombinant
CC protein where the protein is recovered in multimeric form without
CC performing a scaffold refolding step. The protein is useful for treating
CC systemic lupus erythematosus. The present sequence represents a
CC thrombopoietin agonist peptide, which is used to produce fusion proteins
CC of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
RESULT 19
ADM72483
ID ADM72483 standard; peptide; 14 AA.
XX
AC ADM72483;
XX
XX 17-JUN-2004 (first entry)
DT
XX TPO mimetic peptide fragment.
DE
XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
XX WO2004026332-A1.
PN
XX

PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
PT Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.
XX
CC The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
||| ||||| |||||
Db 1 IEGPTLRQWLAARA 14

RESULT 20
ADQ16584
ID ADQ16584 standard; peptide; 14 AA.
XX
AC ADQ16584;
XX
DT 09-SEP-2004 (first entry)

Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
||| ||||| |||||
Db 1 IEGPTLRQWLAARA 14

RESULT 20
ADQ16584
ID ADQ16584 standard; peptide; 14 AA.
XX
AC ADQ16584;
XX
DT 09-SEP-2004 (first entry)

XX Agonist TPO mimetic peptide SEQ ID NO:1.
DE
XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.
XX
OS Unidentified.
XX
PN WO2004050017-A2.
XX
PD 17-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-US036894.
XX
PR 02-DEC-2002; 2002US-00307724.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M;
XX
DR WPI; 2004-460973/43.
XX
PT New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
PS Claim 8; SEQ ID NO 1; 107pp; English.
XX
CC The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents a TPO mimetic peptide.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
||| ||||| |||||
Db 1 IEGPTLRQWLAARA 14

RESULT 21
ADT92482
ID ADT92482 standard; peptide; 14 AA.
XX
AC ADT92482;
XX
DT 13-JAN-2005 (first entry)
XX
DE Thrombopoietin mimetic peptide AF12505 SeqID 5.
XX
KW integrin targeting compound; RGD peptidomimetic; integrin-alpha;
KW angiogenesis; bone metabolism; inflammation; cancer; cell growth;
KW antiangiogenic; osteopathic; antiinflammatory; cytostatic; TPO;
KW thrombopoietin.
XX
OS Unidentified.
XX
PN WO2004091542-A2.
XX
PD 28-OCT-2004.
XX
PF 15-APR-2004; 2004WO-US012034.

XX 15-APR-2003; 2003US-0463456P.
PR 30-SEP-2003; 2003US-0507887P.
XX
XX (COVX-) COVX PHARM INC.
XX
PI Tamiz A, Bradshaw CW;
XX
XX WPI; 2004-784488/77.
XX
XX New nitrogen containing integrin targeting compounds are integrin alpha-v
PT beta-3 antagonists useful for the treatment or prevention of e.g. bone
PT metabolism, inflammation, cancer and cell growth.
XX
XX Example 7; SEQ ID NO 5; 182pp; English.
PS
XX This invention relates to novel nitrogen containing integrin targeting
CC compounds, their stereoisomers, tautomers or salts thereof. Specifically,
CC it refers to integrin targeting compounds that comprise small MW, RGD
CC peptidomimetic integrin targeting agent-linker conjugates that are linked
CC to polymers such as a protein. The present invention describes compounds
CC that can be useful for the treatment or prevention of diseases that
CC involve integrin-alpha, and can be used to treat a defect in
CC angiogenesis, bone metabolism, inflammation, cancer or cell growth.
CC Accordingly, they exhibit antiangiogenic, osteopathic, antiinflammatory
CC and cytostatic activities. This peptide sequence is a TPO mimetic peptide
CC given in an exemplification of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14
RESULT 22
ADU70210
ID ADU70210 standard; peptide; 14 AA.
XX
AC ADU70210;
XX
DT 10-FEB-2005 (first entry)
XX
DE Thrombopoietin receptor modulator peptide #5.
XX
KW hematological disease; drug delivery; thrombopoietin receptor modulator;
KW diagnosis; pharmaceutical.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= OTHER
FT /note= "Covalently linked to a spacer/PEG moiety which is
FT in turn lionked to another copy of this sequence forming
FT a dimer"
FT Modified-site 14 /note= "Ala is amidated"
FT
FT WO2004100997-A2.
XX
PN 25-NOV-2004.
XX
PD 12-MAY-2004; 2004WO-US014887.
XX
PF 12-MAY-2003; 2003US-0469996P.
XX
PR (AFFY-) AFFYMAX INC.
XX

PI Holmes C, Yin K, Tumelty D;
XX WPI; 2005-039305/04.
XX
PT New peptide-based compound comprising a peptide moiety, a spacer moiety
PT and a water-soluble polymer moiety useful as pharmaceuticals.
XX
XX Example 15; Page 46; 57pp; English.
PS
XX The invention relates to a peptide-based compound (I) comprising a
CC peptide moiety, a spacer moiety and a water-soluble polymer moiety (e.g.
CC polyethylene glycol, PEG). Also included is a pharmaceutical composition
CC comprising (I) and one or more pharmaceutically acceptable diluents,
CC preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. The
CC peptide moiety comprises one or more peptides that bind to erythropoietin
CC -receptors or thrombopoietin-receptors. The peptide based compound is
CC useful as a pharmaceutical. The peptide moiety particularly
CC erythropoietin receptors are essential in the process of red blood cell
CC formation and has useful applications in both diagnosis and the treatment
CC of blood disorders characterized by low or defective red blood cell
CC production. The peptide based compound has improved antigenicity and
CC durability. Polyethylene glycolylation leads to improved drug efficacy and
CC safety, increases bioavailability and decreases immunogenicity and dosing
CC frequency. The present sequence is a thrombopoietin-receptor modulating
CC peptide.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14
RESULT 23
ADU75982
ID ADU75982 standard; peptide; 14 AA.
XX
AC ADU75982;
XX
DT 10-FEB-2005 (first entry)
XX
DE Peptide-based compound associated peptide #3.
XX
KW antianemic; pharmaceutical; blood; blood disorder.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= OTHER
FT /note= "OTHER= Two of these molecules are optionally
FT attached by the N-terminal by a spacer molecule that can
FT undergo PEGylation"
FT
FT WO2004101600-A2.
XX
PN 25-NOV-2004.
XX
PD 12-MAY-2004; 2004WO-US014888.
XX
PF 12-MAY-2003; 2003US-0470246P.
XX
PR (AFFY-) AFFYMAX INC.
XX
PA Yin K, Holmes C, Lalonde G, Balu P, Tumelty D, Schatz P;
XX WPI; 2005-039327/04.
XX
PT Peptide-based compound useful as e.g. pharmaceutical for the treatment of

PT blood disorders comprises a peptide moiety and a linear
PT poly(ethyleneglycol) moiety.
XX
PS Disclosure; Page 47; 56pp; English.
XX
CC The invention describes a peptide-based compound (A) comprising a peptide
CC moiety (B) and a linear poly(ethyleneglycol) moiety (C) having molecular
CC weight of more than 20 kDaltons. The compounds are useful as
CC pharmaceuticals for the treatment of blood disorders characterized by low
CC or defective red blood cell production, for reducing immunogenicity and
CC as peptide modifying reagent. The compounds exhibit improved antigenicity
CC and durability. This is the amino acid sequence of peptide that may be
CC present in the peptide moieties of the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 24
ADV44319
ID ADV44319 standard; peptide; 14 AA.
XX
AC ADV44319;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 1.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
XX WPI; 2005-031588/03.
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Claim 8; SEQ ID NO 1; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an

CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 25
AEB12792
ID AEB12792 standard; peptide; 14 AA.
XX
AC AEB12792;
XX
DT 08-SEP-2005 (first entry)
XX
DE TPO mimetic peptide, SEQ ID 1.
XX
KW TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.
XX
OS Synthetic.
XX
PN WO2005060642-A2.
XX
PD 07-JUL-2005.
XX
PF 15-DEC-2004; 2004WO-US041946.
XX
PR 15-DEC-2003; 2003US-00737290.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
XX WPI; 2005-479402/48.
XX
PT New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
PS Example 1; SEQ ID NO 1; 152pp; English.
XX
CC The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under
CC conditions suitable for expression of the immunoglobulin or its
CC fragment), a composition comprising an immunoglobulin (or its fragment)
CC and a pharmaceutically acceptable carrier, treating congestive heart

CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
CC glucagons, glucagons mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1),
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to
CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a Thrombopoietin (TPO) mimetic , erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 26
AAW35416
ID AAW35416 standard; peptide; 15 AA.

XX AAW35416;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

XX Synthetic.

XX Key Location/Qualifiers

FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical
FT peptide"

FT Modified-site 15 /note= "NH2-Ala"

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US0009623.

XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX (GLAX) GLAXO GROUP LTD.
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 73; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines

XX Sequence 15 AA;

Query Match 100.0%; Score 73; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

RESULT 27
AAW36776

ID AAW36776 standard; peptide; 15 AA.

XX AAW36776;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

XX Synthetic.

XX Key Location/Qualifiers

FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical
FT peptide"

FT Modified-site 15 /note= "NH2-Ala"

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US0009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
DR Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX Example 9; Page 77; 106pp; English.
PS The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
Db 2 IEGPTLRQWLAARA 15
| | | | | | | | | | | | | | | |
RESULT 28
AAW66712
ID AAW66712 standard; peptide; 15 AA.
XX
AC AAW66712;
XX
DT 01-DEC-1998 (first entry)
XX
DE Peptide chain of compound which binds to the thrombopoietin receptor.
XX
KW thrombopoietin receptor; haematological disorder; screening; agonist;
KW assay; megakaryocyte; blood disorder; thrombocytopenia; TPO.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..14
FT /note= "thrombopoietin receptor agonist"
FT Modified-site 15
FT /note= "Epsilon amino group of Lys, in its amide form, is
FT attached to another peptide chain identical to the region
FT (residues 1 to 14) of this peptide"
XX
PN WO9825965-A2.
XX
PD 18-JUN-1998.
XX
PF 09-DEC-1997; 97WO-EP006850.
XX
PR 11-DEC-1996; 96US-00764640.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 1998-377261/32.
XX
DR New peptide compound(s) which can bind and activate thrombopoietin
PT receptor - may be used in treating haematological disorders and in
PT methods for screening for new thrombopoietin receptor agonists.
PT

PS Claim 2; Page 60; 78pp; English.
XX
CC The invention relates to peptide compounds composed of two peptide chains
CC attached to each of the amino groups of a single Lys in the amide form.
CC The compounds are of formula (Pep1) (Pep2)K(NH2), where Pep1 is of
CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9-X10; and pep2 is of
CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9'-X10'. X1 = H or acyl; X2
CC = Gly or Sar (sarcosine); X3 = Arg, Ala, Nle (norleucine) or N-
CC acetyllysine; X4 = Gln or Glu; X5 = Trp, L-1-naphthylalanine or Phe; X6 =
CC Ala, 5-aminopentanoic acid or 2-aminobutyric acid; X7 = Ala,
CC diphenylalanine, or is absent; X8 = Arg, p- amino-phenylalanine, N-
CC acetyl-lysine, or is absent; X9, X9' = Ala, beta Ala, N-methyl-alanine,
CC Sar, or is absent; X10, X10' = beta Ala or is absent. The new peptides
CC are capable of binding to, and activating, the thrombopoietin (TPO)
CC receptor. They may be used in vitro as tools for understanding the
CC biological role of TPO. They may be used as competitive binders in assays
CC to screen for new TPO receptor agonists. They may be used as reagents for
CC detecting TPO receptors in living cells, biological fluids, etc. They may
CC be used to maintain growth and proliferation of TPO-dependent cells and
CC for in vitro expansion of megakaryocytes. They may be used to activate
CC TPO receptors in vivo, e.g., to treat blood disorders or
CC thrombocytopenia associated with bone marrow transfusions, radiotherapy
CC or chemotherapy. The present sequence represents a specific example of
CC (Pep1)K(NH2)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | | |
RESULT 29
AAB20684
ID AAB20684 standard; peptide; 15 AA.
XX
AC AAB20684;
XX
DT 20-DEC-2000 (first entry)
XX
DE Thrombocyte generation promoting peptide.
XX
KW Thrombocyte; promotion; generation.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 15
FT /note= "optionally amidated; optionally attached to the C
FT -terminal cysteine of a similar peptide"
XX
PN CN1254718-A.
XX
PD 31-MAY-2000.
XX
PF 20-NOV-1998; 98CN-00125011.
XX
PR 20-NOV-1998; 98CN-00125011.
XX
PA (BIOL-) INST BIOLOGICAL ENG CHINESE ACAD MILITAR.
XX
PI Cheng D, Li C, Huang P;
XX
DR WPI; 2000-533568/49.
XX
PT Active peptide.
XX
PS Claim 1; Page 1; 5pp; Chinese.
XX

CC The present invention discloses an active peptide which promotes
CC thrombocyte generation. The active peptide can be synthesised by a
CC polypeptide solid-phase synthesis method, and has the monomer sequence of
CC IEGPTLRQWLAARAC and the amidated peptide chain structure of
CC IEGPTLRQWLAARAC-NH2. Its activity is increased by 20 times for its
CC monomer, or by 10 times for the amidated peptide chain compared with the
CC monomer, or by 100 times for its dimer compared with its monomer
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 30
AAU25996
ID AAU25996 standard; peptide; 15 AA.
XX
AC AAU25996;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #182.
XX

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN U96251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX

PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX

XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX

PS Disclosure; Col 143-144; 128pp; English.
XX

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

RESULT 31
AAU25831
ID AAU25831 standard; peptide; 15 AA.
XX
AC AAU25831;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #17.
XX

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX

PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX

XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX

PS Claim 1; Col 69-70; 128pp; English.
XX

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro

Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
 |||||
DB 2 IEGPTLRQWLAARA 15

RESULT 34
ADM72485

ID ADM72485 standard; peptide; 15 AA.
XX
AC ADM72485;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX

Key Location/Qualifiers
FT Modified-site 15 /label= bAla
FT /note= "beta-alanine"
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX

Increasing hematopoietic stem cell production in subject, useful in
reducing the incidence of delayed primary engraftment, comprises
administering a Thrombopoietin mimetic compound e.g., a peptide to a
subject.

Disclosure; Fig 2; 32pp; English.

The invention relates to a method (M1) for increasing haematopoietic stem
cell production in a subject which involves administering a
Thrombopoietin (TPO) mimetic compound to the subject. Also included is
another method (M2) of providing haematopoietic stem cells to a subject
which involves administering a TPO mimetic compound to a subject to
enhance expansion of a stem cell population within bone marrow and/or
mobilize stem cells in peripheral circulation, harvesting one or more of
the bone marrow stem cells or the stem cells in the peripheral
circulation, and transplanting the harvested stem cells into the subject.
A method (M3) is also provided for reducing a time to engraftment
following reinfusion of stem cells in a subject, involves administering a
TPO mimetic compound to the subject, enhancing the expansion of the stem
cell population within bone marrow and/or mobilizing the stem cells in
peripheral circulation, harvesting one or more of the bone marrow stem
cells or one or more of the stem cells in the peripheral circulation, and
transplanting the one or more harvested stem cells into the subject. TPO
mimetic compounds are disclosed as peptides, including cyclic or modified
peptides. (M1) is useful for increasing haematopoietic stem cell
production in a subject e.g., human. (M3) is useful for reducing time to
engraftment following reinfusion of stem cells, reducing the incidence of
delayed primary engraftment, reducing the incidence of secondary failure
of platelet production and reducing the time of platelet and/or
neutrophil engraftment following reinfusion of stem cells in a subject.

(M1) is also useful for increasing the number of stem cells from a donor
whose cells are then used for rescue of recipient subject. Also useful in
the treatment of thrombocytopenia. (M1) enables transplantation to
proceed in patients who would not otherwise be considered as candidates
because of unacceptably high risk of failed engraftment, reduces the
number of aphereses required to generate a minimum acceptable harvest,
reduces the incidence of primary and secondary failure of engraftment by
increasing the number of haematopoietic stem cells (HSCs) available for
transplantation and reduces the time required for primary engraftment.
The present sequence represents an example of TPO mimetic peptide
fragment.

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
 |||||
DB 1 IEGPTLRQWLAARA 14

RESULT 35
ADM72479

ID ADM72479 standard; peptide; 15 AA.
XX
AC ADM72479;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX

Key Location/Qualifiers
FT Modified-site 15 /note= "beta-alanine"
FT
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX

Increasing hematopoietic stem cell production in subject, useful in
reducing the incidence of delayed primary engraftment, comprises
administering a Thrombopoietin mimetic compound e.g., a peptide to a
subject.

Disclosure; Fig 2; 32pp; English.

The invention relates to a method (M1) for increasing haematopoietic stem
cell production in a subject which involves administering a
Thrombopoietin (TPO) mimetic compound to the subject. Also included is
another method (M2) of providing haematopoietic stem cells to a subject
which involves administering a TPO mimetic compound to a subject to
enhance expansion of a stem cell population within bone marrow and/or
mobilize stem cells in peripheral circulation, harvesting one or more of
the bone marrow stem cells or the stem cells in the peripheral
circulation, and transplanting the harvested stem cells into the subject.
A method (M3) is also provided for reducing a time to engraftment
following reinfusion of stem cells in a subject, involves administering a
TPO mimetic compound to the subject, enhancing the expansion of the stem
cell population within bone marrow and/or mobilizing the stem cells in
peripheral circulation, harvesting one or more of the bone marrow stem
cells or one or more of the stem cells in the peripheral circulation, and
transplanting the one or more harvested stem cells into the subject. TPO
mimetic compounds are disclosed as peptides, including cyclic or modified
peptides. (M1) is useful for increasing haematopoietic stem cell
production in a subject e.g., human. (M3) is useful for reducing time to
engraftment following reinfusion of stem cells, reducing the incidence of
delayed primary engraftment, reducing the incidence of secondary failure
of platelet production and reducing the time of platelet and/or
neutrophil engraftment following reinfusion of stem cells in a subject.

CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the incidence of secondary failure
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

SQ Sequence 15 AA;
Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
|||||
Db 1 IEGPTLRQWLAARA 14

RESULT 36
ADM72478
ID ADM72478 standard; peptide; 15 AA.

XX AC ADM72478;

DT 17-JUN-2004 (first entry)

XX TPO mimetic peptide fragment.

DE TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX Key Location/Qualifiers
FH Modified-site 15

FT /note= "Lys (15) is linked to one copy of the TPO mimetic
FT peptide through the alpha amino group and to a second
FT copy of the peptide (not shown) via the omega amino
FT group"

XX WO2004026332-A1.

XX PD 01-APR-2004.

XX PF 18-SEP-2003; 2003WO-US029701.

XX PR 18-SEP-2002; 2002US-0411700P.

XX PR 18-SEP-2002; 2002US-0411779P.

XX (THRE-) 3-DIMENSIONAL PHARM INC.

XX Kaushansky K, Macdonald BR;

XX WPI; 2004-283153/26.

XX Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.

XX Disclosure; Fig 2; 32pp; English.

XX The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the time of platelet and/or
CC of platelet production and reducing the time of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
|||||
Db 1 IEGPTLRQWLAARA 14

RESULT 37

ADM72533

ID ADM72533 standard; peptide; 15 AA.

XX AC ADM72533;

XX 17-JUN-2004 (first entry)

XX TPO mimetic peptide fragment.

DE TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX Key Location/Qualifiers

FH Disulfide-bond 1

FT /note = connected to a cysteine residue of a similar TPO
FT mimetic peptide (not shown) by a disulphide bridge

CC fragment.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14

RESULT 39
ADM72523
ID ADM72523 standard; peptide; 15 AA.
XX
AC ADM72523;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 15
FT /note= "beta-alanine"
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
PT Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.
XX
CC The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to

CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the incidence of secondary failure
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment, by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
Db 1 IEGPTLRQWLAARA 14

RESULT 40
ADM72482
ID ADM72482 standard; peptide; 15 AA.
XX
AC ADM72482;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 15
FT /note= "Lys (15) is linked to one copy of the TPO mimetic
FT peptide through the alpha amino group and to a second
FT copy of the peptide (not shown) via the omega amino
FT group"
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
PT Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.
XX
CC The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a

CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the incidence of secondary failure
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 41
ADQ16585
ID ADQ16585 standard; peptide; 15 AA.

XX AC ADQ16585;

DT 09-SEP-2004 (first entry)

XX TPO mimetic peptide SEQ ID NO:2.

DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.

XX OS Unidentified.

XX PN WO2004050017-A2.

XX 17-JUN-2004.

PF 17-NOV-2003; 2003WO-US036894.

XX 02-DEC-2002; 2002US-00307724.

XX (ALEX-) ALEXION PHARM INC.

XX PA Bowdish KS, Frederickson S, Renshaw M;

XX WPI; 2004-460973/43.

XX
PT New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.

XX PS Disclosure; SEQ ID NO 2; 107pp; English.

XX
CC The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents a TPO mimetic peptide.

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 42
ADT92483

ID ADT92483 standard; peptide; 15 AA.

XX AC ADT92483;

XX 13-JAN-2005 (first entry)

DE Modified thrombopoietin mimetic peptide AF12505 with amino Cys residue.

XX
KW integrin targeting compound; RGD peptidomimetic; integrin-alpha;
KW angiogenesis; bone metabolism; inflammation; cancer; cell growth;
KW antiangiogenic; osteopathic; antiinflammatory; cytostatic; TPO;
KW thrombopoietin.

XX OS Synthetic.

OS Unidentified.

XX PN WO2004091542-A2.

XX 28-OCT-2004.

XX 15-APR-2004; 2004WO-US012034.

XX 15-APR-2003; 2003US-0463456P.

PR 30-SEP-2003; 2003US-0507887P.

XX (COVX-) COVX PHARM INC.

XX Tamiz A, Bradshaw CW;

XX WPI; 2004-784488/77.

XX
PT New nitrogen containing integrin targeting compounds are integrin alpha-v
PT beta-3 antagonists useful for the treatment or prevention of e.g. bone
PT metabolism, inflammation, cancer and cell growth.

XX PS Example 7; SEQ ID NO 6; 182pp; English.

XX
CC This invention relates to novel nitrogen containing integrin targeting
CC compounds, their stereoisomers, tautomers or salts thereof. Specifically,
CC it refers to integrin targeting compounds that comprise small MW, RGD
CC peptidomimetic integrin targeting agent-linker conjugates that are linked
CC to polymers such as a protein. The present invention describes compounds

CC that can be useful for the treatment or prevention of diseases that
CC involve integrin-alpha, and can be used to treat a defect in
CC angiogenesis, bone metabolism, inflammation, cancer or cell growth.
CC Accordingly, they exhibit antiangiogenic, osteopathic, antiinflammatory
CC and cytostatic activities. This peptide sequence is a modified TPO
CC mimetic peptide with amino Cys given in an exemplification of the
CC invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

RESULT 43
ADU70209
ID ADU70209 standard; peptide; 15 AA.
XX
AC ADU70209;
XX
DT 10-FEB-2005 (first entry)
XX
DE Thrombopoietin receptor modulator peptide #4.
XX
KW hematological disease; drug delivery; thrombopoietin receptor modulator;
KW diagnosis; pharmaceutical.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Ile is acetylated"
FT Modified-site 15 /label= OTHER
FT /note= "Covalently linked to a spacer/PEG moiety which is
FT in turn lionked to another copy of this sequence forming
FT a dimer"
XX
PN WO2004100997-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014887.
XX
PR 12-MAY-2003; 2003US-0469996P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Holmes C, Yin K, Tumelty D;
XX
DR WPI; 2005-039305/04.
XX
PT New peptide-based compound comprising a peptide moiety, a spacer moiety
PT and a water-soluble polymer moiety useful as pharmaceuticals.
XX
PS Example 14; Page 43; 57pp; English.
XX
CC The invention relates to a peptide-based compound (I) comprising a
CC peptide moiety, a spacer moiety and a water-soluble polymer moiety (e.g.
CC polyethylene glycol, PEG). Also included is a pharmaceutical composition
CC comprising (I) and one or more pharmaceutically acceptable diluents, The
CC preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. The
CC peptide moiety comprises one or more peptides that bind to erythropoietin
CC -receptors or thrombopoietin-receptors. The peptide based compound is
CC useful as a pharmaceutical. The peptide moiety particularly
CC erythropoietin receptors are essential in the process of red blood cell
CC formation and has useful applications in both diagnosis and the treatment
CC of blood disorders characterized by low or defective red blood cell

CC production. The peptide based compound has improved antigenicity and
CC durability. Polyethylene glycolylation leads to improved drug efficacy and
CC safety, increases bioavailability and decreases immunogenicity and dosing
CC frequency. The present sequence is a thrombopoietin-receptor modulating
CC peptide.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 44
ADU75981
ID ADU75981 standard; peptide; 15 AA.
XX
AC ADU75981;
XX
DT 10-FEB-2005 (first entry)
XX
DE Peptide-based compound associated peptide #2.
XX
KW antianemic; pharmaceutical; blood; blood disorder.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "N-terminal acetyl"
FT Modified-site 15 /label= OTHER
FT /note= "OTHER= Two of these molecules are optionally
FT attached by the C-terminal by a spacer molecule that can
FT undergo PEGylation"
XX
PN WO2004101600-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014888.
XX
PR 12-MAY-2003; 2003US-0470246P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Yin K, Holmes C, Lalonde G, Balu P, Tumelty D, Schatz P;
XX
DR WPI; 2005-039327/04.
XX
PT Peptide-based compound useful as e.g. pharmaceutical for the treatment of
PT blood disorders comprises a peptide moiety and a linear
PT poly(ethyleneglycol) moiety.
XX
PS Disclosure; Page 44; 56pp; English.
XX
CC The invention describes a peptide-based compound (A) comprising a peptide
CC moiety (B) and a linear poly(ethyleneglycol) moiety (C) having molecular
CC weight of more than 20 KDaltons. The compounds are useful as
CC pharmaceuticals for the treatment of blood disorders characterized by low
CC or defective red blood cell production, for reducing immunogenicity and
CC as peptide modifying reagent. The compounds exhibit improved antigenicity
CC and durability. This is the amino acid sequence of peptide that may be
CC present in the peptide moieties of the invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 45
ADV44320
ID ADV44320 standard; peptide; 15 AA.
XX
AC ADV44320;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 2.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-031588/03.
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Disclosure; SEQ ID NO 2; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 46
AEB12793
ID AEB12793 standard; peptide; 15 AA.
XX
AC AEB12793;
XX
DT 08-SEP-2005 (first entry)
XX
DE TPO mimetic peptide, SEQ ID 2.
XX
KW TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.
XX
OS Synthetic.
XX
PN WO2005060642-A2.
XX
PD 07-JUL-2005.
XX
PF 15-DEC-2004; 2004WO-US041946.
XX
PR 15-DEC-2003; 2003US-00737290.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-479402/48.
XX
PT New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
PS Disclosure; SEQ ID NO 2; 152pp; English.
XX
CC The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under
CC conditions suitable for expression of the immunoglobulin or its
CC fragment), a composition comprising an immunoglobulin (or its fragment)
CC and a pharmaceutically acceptable carrier, treating congestive heart
CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
CC glucagons, glucagons mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1),
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to

CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a Thrombopoietin (TPO) mimetic , erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 47
AAW19534
ID AAW19534 standard; protein; 16 AA.
XX
AC AAW19534;

XX 10-SEP-1997 (first entry)
DT
XX Thrombopoietin receptor binding compound peptide (part of a dimer).
DE
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
KW
XX Synthetic.

XX Key Location/Qualifiers
FT Modified-site 15 /label= bAla
FT Cross-links 16
FT Modified-site 16 /note= "Linked to the omega Ala in AAW09468"
FT
FT /note= "In amide form"

XX WO9640189-A1.
PN
XX 19-DEC-1996.
PD
XX 05-JUN-1996; 96WO-US008998.
PF
XX 07-JUN-1995; 95US-00472371.
PR
XX 07-JUN-1995; 95US-00473604.
PR
XX 07-JUN-1995; 95US-00476168.
PR
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00484090.
PR
XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
PI
XX WPI; 1997-051883/05.
DR

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.

PS Claim 30; Page 91; 106pp; English.
XX The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It is part of a dimer linked by the omega amino acid to
CC the omega amino acid in the sequence in AAW09468. The compound can be
CC used for treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 48
AAW33035
ID AAW33035 standard; peptide; 16 AA.
XX

AC AAW33035;
XX 11-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
DE
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.

XX Key Location/Qualifiers
FT Cross-links 14 /note= "epsilon amino group of Lys16 linked to terminal
FT carboxy group of AAW33034"
FT Modified-site 15 /label= bAla
FT
XX WO9640750-A1.
PN
XX 19-DEC-1996.
PD
XX 07-JUN-1996; 96WO-US009623.
PF
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
PI
XX WPI; 1997-052226/05.
DR
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 30; Page 91; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia

CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 49
AAW36775
ID AAW36775 standard; peptide; 16 AA.
XX
AC AAW36775;
XX 11-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.

DE
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT Disulfide-bond 1. .16
FT Modified-site 16
FT /note= "NH2-Cys"
XX

PN WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 77; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

RESULT 50
AAW36771
ID AAW36771 standard; peptide; 16 AA.
XX
AC AAW36771;
XX 11-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT Disulfide-bond 1. .16
FT Modified-site 16
FT /note= "NH2-Cys"
XX

PN WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 76; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

Search completed: May 12, 2006, 10:36:51

Job time : 105.809 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16.9915 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-13
Perfect score: 73
Sequence: 1 IEGPTLRQWLAARA 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	49.5	67.8	333	2 A36925	transcription acti
2	47	64.4	296	2 AG0147	probable membrane
3	46	63.0	306	2 D70601	UTP-glucose-1-phos
4	44	60.3	200	2 T23485	hypothetical prote
5	44	60.3	207	2 T37464	probable glutathio
6	44	60.3	536	1 SYECBB	2,3-dihydroxybenzo
7	44	60.3	536	2 E85558	2,3-dihydroxybenzo
8	44	60.3	536	2 A99708	2,3-dihydroxybenzo
9	43	58.9	285	2 G71337	probable dimethyla
10	43	58.9	683	2 B71325	conserved hypothet
11	42	57.5	473	2 E84853	hypothetical prote
12	42	57.5	1019	2 T11560	pol polyprotein -
13	41	56.2	195	2 F91171	probable phosphopa
14	41	56.2	195	2 F86017	probable phosphopa
15	41	56.2	195	2 S47694	hypothetical 21.8K
16	41	56.2	249	2 E87575	ABC transporter, A
17	41	56.2	306	2 T45453	UTP-glucose-1-phos
18	41	56.2	326	2 C24430	glyceraldehyde-3-p
19	41	56.2	336	1 DEPZG	glyceraldehyde-3-p
20	41	56.2	337	2 A35080	glyceraldehyde-3-p
21	41	56.2	338	1 DEIS3C	glyceraldehyde-3-p
22	41	56.2	338	2 JQ1287	glyceraldehyde-3-p
23	41	56.2	719	2 B95325	conserved hypothet
24	41	56.2	750	2 A97501	topoisomerase iv c
25	41	56.2	750	2 AE2719	topoisomerase IV s
26	40	54.8	239	2 S25204	smx protein - Str
27	40	54.8	463	2 S27491	hypothetical prote
28	40	54.8	530	2 A81958	probable permease
29	40	54.8	531	2 E81015	ABC transporter, p

30	40	54.8	656	2 S30484	pol polyprotein -
31	40	54.8	656	2 S30483	pol polyprotein -
32	40	54.8	721	2 A39707	erythrocyte membra
33	40	54.8	1123	2 T51517	telomerase reverse
34	40	54.8	1712	1 CGHU2B	collagen alpha 2(I
35	39.5	54.1	325	2 A84326	hypothetical prote
36	39	53.4	131	2 S74539	hypothetical prote
37	39	53.4	217	2 S46354	pol polyprotein -
38	39	53.4	267	2 I40327	baf protein - Bord
39	39	53.4	331	2 B48445	glyceraldehyde-3-p
40	39	53.4	331	2 A72514	hypothetical prote
41	39	53.4	400	2 C87021	serine-threonine p
42	39	53.4	600	2 C83221	transport protein
43	39	53.4	791	2 A82291	c-di-GMP phosphodi
44	39	53.4	1034	1 GNLJCA	HIV-1 retropepsin
45	39	53.4	1035	1 GNLJGG	HIV-1 retropepsin
46	39	53.4	1036	1 GNLJG2	HIV-1 retropepsin
47	39	53.4	1055	1 GNLJST	HIV-1 retropepsin
48	39	53.4	1055	2 S53092	pol polyprotein -
49	39	53.4	1058	2 S08436	pol polyprotein -
50	39	53.4	3345	2 T13423	hypothetical prote
51	38	52.1	134	2 B75468	hypothetical prote
52	38	52.1	197	2 G82973	transcription regu
53	38	52.1	246	2 AH0190	probable oxidoredu
54	38	52.1	247	2 PQ0178	glyceraldehyde-3-p
55	38	52.1	295	2 T07730	glyceraldehyde-3-p
56	38	52.1	297	2 B87109	integrase/recombin
57	38	52.1	311	1 RGECK	regulatory protein
58	38	52.1	311	2 AH0867	transcription acti
59	38	52.1	311	2 C85936	positive regulator
60	38	52.1	311	2 H91090	positive regulator
61	38	52.1	314	2 H70723	hypothetical prote
62	38	52.1	337	1 DEPJG	glyceraldehyde-3-p
63	38	52.1	337	1 DESKG	glyceraldehyde-3-p
64	38	52.1	337	1 DEUSGM	glyceraldehyde-3-p
65	38	52.1	337	1 DEZMGC	glyceraldehyde-3-p
66	38	52.1	339	2 A83358	hypothetical prote
67	38	52.1	350	2 B38535	A/G-specific adeni
68	38	52.1	350	2 H85953	adenine glycosylas
69	38	52.1	350	2 E91108	adenine glycosylas
70	38	52.1	360	2 S38570	glyceraldehyde-3-p
71	38	52.1	469	2 AD1926	hypothetical prote
72	38	52.1	589	2 F87626	sensor histidine k
73	38	52.1	635	2 A87433	hypothetical prote
74	38	52.1	816	2 A71006	probable polA prot
75	38	52.1	904	2 C70559	adenylate cyclase
76	38	52.1	1155	2 AC2426	transposase - Cory
77	37.5	51.4	436	2 JC4742	HIV-1 retropepsin
78	37	50.7	151	2 S63748	hypothetical prote
79	37	50.7	165	2 F87542	glyceraldehyde-3-p
80	37	50.7	234	2 PQ0179	glyceraldehyde-3-p
81	37	50.7	305	2 A24159	glyceraldehyde-3-p
82	37	50.7	335	2 S29813	glyceraldehyde-3-p
83	37	50.7	337	1 DEBHG	glyceraldehyde-3-p
84	37	50.7	337	2 S42479	glyceraldehyde-3-p
85	37	50.7	337	2 T02723	glyceraldehyde-3-p
86	37	50.7	337	2 T02722	glyceraldehyde-3-p
87	37	50.7	338	1 DENDG	glyceraldehyde-3-p
88	37	50.7	338	2 T06781	glyceraldehyde-3-p
89	37	50.7	341	1 DEJMG	glyceraldehyde-3-p
90	37	50.7	341	2 T08147	glyceraldehyde-3-p
91	37	50.7	341	2 AG0195	probable exported
92	37	50.7	352	2 G83636	conserved hypothet
93	37	50.7	391	2 T36739	hypothetical prote
94	37	50.7	407	2 A86298	hypothetical prote
95	37	50.7	422	2 F96826	hypothetical prote
96	37	50.7	433	2 S51837	glyceraldehyde-3-p
97	37	50.7	433	2 S51836	glyceraldehyde-3-p
98	37	50.7	438	2 G87337	membrane protein,
99	37	50.7	480	2 H84747	probable steroid d
100	37	50.7	486	2 B86411	protein F3M18.4 [i
101	37	50.7	544	2 A72459	probable glutamyl-
102	37	50.7	547	2 T36550	hypothetical prote

249	34	46.6	98	2	A70301	ribosomal protein	322	34	46.6	219	2	S32056	RNA-directed DNA p
250	34	46.6	114	2	AI0294	probable membrane	323	34	46.6	219	2	S32059	RNA-directed DNA p
251	34	46.6	119	2	S75551	hypothetical prote	324	34	46.6	219	2	S32131	RNA-directed DNA p
252	34	46.6	129	2	S63743	HIV-1 retropepsin	325	34	46.6	219	2	S32071	RNA-directed DNA p
253	34	46.6	130	2	S63754	HIV-1 retropepsin	326	34	46.6	219	2	S32127	RNA-directed DNA p
254	34	46.6	132	2	S63733	HIV-1 retropepsin	327	34	46.6	219	2	S32134	RNA-directed DNA p
255	34	46.6	136	2	S63746	HIV-1 retropepsin	328	34	46.6	219	2	S32093	RNA-directed DNA p
256	34	46.6	140	2	S63737	HIV-1 retropepsin	329	34	46.6	219	2	S32075	RNA-directed DNA p
257	34	46.6	141	2	S63750	HIV-1 retropepsin	330	34	46.6	219	2	S32079	RNA-directed DNA p
258	34	46.6	142	2	AC0928	probable regulator	331	34	46.6	219	2	S32119	RNA-directed DNA p
259	34	46.6	143	2	S63735	HIV-1 retropepsin	332	34	46.6	219	2	S32138	RNA-directed DNA p
260	34	46.6	145	2	S63745	HIV-1 retropepsin	333	34	46.6	219	2	S32129	RNA-directed DNA p
261	34	46.6	145	2	S07957	hypothetical prote	334	34	46.6	219	2	S32097	RNA-directed DNA p
262	34	46.6	149	2	S63734	HIV-1 retropepsin	335	34	46.6	219	2	S32117	RNA-directed DNA p
263	34	46.6	151	2	S63731	HIV-1 retropepsin	336	34	46.6	219	2	S32077	RNA-directed DNA p
264	34	46.6	153	2	A97524	hypothetical prote	337	34	46.6	219	2	S32098	RNA-directed DNA p
265	34	46.6	158	2	AI0605	conserved hypothet	338	34	46.6	219	2	S32051	RNA-directed DNA p
266	34	46.6	158	2	E64823	hypothetical prote	339	34	46.6	219	2	S32132	RNA-directed DNA p
267	34	46.6	158	2	E90745	probable sensory t	340	34	46.6	219	2	S32058	RNA-directed DNA p
268	34	46.6	160	2	S63753	HIV-1 retropepsin	341	34	46.6	219	2	S32140	RNA-directed DNA p
269	34	46.6	162	2	S63732	HIV-1 retropepsin	342	34	46.6	223	2	S44974	lmbU protein - Str
270	34	46.6	174	2	JCL153	hypothetical 19.7K	343	34	46.6	244	2	D86275	F7A19.26 protein -
271	34	46.6	176	2	S63747	HIV-1 retropepsin	344	34	46.6	254	2	F87652	hypothetical prote
272	34	46.6	178	2	AB2743	hypothetical prote	345	34	46.6	254	2	C95273	hypothetical prote
273	34	46.6	180	2	T09063	hypothetical prote	346	34	46.6	255	2	A45881	MHC class II histo
274	34	46.6	184	2	H83409	hypothetical prote	347	34	46.6	255	2	AB3572	succinoglycan bios
275	34	46.6	193	2	A53835	SEC63 protein comp	348	34	46.6	274	2	T05509	hypothetical prote
276	34	46.6	197	2	AE2063	hypothetical prote	349	34	46.6	277	2	I51117	proopiomelanocorti
277	34	46.6	206	2	S38626	glutathione transf	350	34	46.6	281	2	H97548	(ac007190) f23n19.
278	34	46.6	215	2	D75379	hypothetical prote	351	34	46.6	281	2	AF2768	phytoene synthase
279	34	46.6	219	2	S32096	RNA-directed DNA p	352	34	46.6	289	2	JQ0059	hypothetical 31.6K
280	34	46.6	219	2	S32157	RNA-directed DNA p	353	34	46.6	295	2	S76136	hypothetical prote
281	34	46.6	219	2	S32080	RNA-directed DNA p	354	34	46.6	298	2	S44994	class I histocompa
282	34	46.6	219	2	S32062	RNA-directed DNA p	355	34	46.6	302	2	E75414	pecM-related prote
283	34	46.6	219	2	S32047	RNA-directed DNA p	356	34	46.6	313	2	AC0098	transcription acti
284	34	46.6	219	2	S32065	RNA-directed DNA p	357	34	46.6	315	2	G95415	probable lysR-fami
285	34	46.6	219	2	S32122	RNA-directed DNA p	358	34	46.6	317	2	F83172	probable D-amino a
286	34	46.6	219	2	S32060	RNA-directed DNA p	359	34	46.6	320	2	T35265	probable D-amino a
287	34	46.6	219	2	S32118	RNA-directed DNA p	360	34	46.6	321	2	C70653	probable prephenat
288	34	46.6	219	2	S32057	RNA-directed DNA p	361	34	46.6	322	2	F86918	probable prephenat
289	34	46.6	219	2	S32048	RNA-directed DNA p	362	34	46.6	324	2	F95306	probable glutamate
290	34	46.6	219	2	S32094	RNA-directed DNA p	363	34	46.6	331	2	F82131	glyceraldehyde 3-p
291	34	46.6	219	2	S32061	RNA-directed DNA p	364	34	46.6	332	2	T44437	aminodeoxychorism
292	34	46.6	219	2	S32084	RNA-directed DNA p	365	34	46.6	333	1	DELOG3	glyceraldehyde-3-p
293	34	46.6	219	2	S32152	RNA-directed DNA p	366	34	46.6	336	1	DEASG3	glyceraldehyde-3-p
294	34	46.6	219	2	S32089	RNA-directed DNA p	367	34	46.6	337	1	DEJJGC	glyceraldehyde-3-p
295	34	46.6	219	2	S32073	RNA-directed DNA p	368	34	46.6	337	1	DEYDGC	glyceraldehyde-3-p
296	34	46.6	219	2	S32049	RNA-directed DNA p	369	34	46.6	337	2	S29814	glyceraldehyde-3-p
297	34	46.6	219	2	S32120	RNA-directed DNA p	370	34	46.6	337	2	S26946	glyceraldehyde-3-p
298	34	46.6	219	2	S32086	RNA-directed DNA p	371	34	46.6	337	2	S26863	glyceraldehyde-3-p
299	34	46.6	219	2	S32133	RNA-directed DNA p	372	34	46.6	338	2	T47218	glyceraldehyde-3-p
300	34	46.6	219	2	S32081	RNA-directed DNA p	373	34	46.6	339	2	G64041	glyceraldehyde-3-p
301	34	46.6	219	2	S32139	RNA-directed DNA p	374	34	46.6	342	1	S64042	porphobilinogen sy
302	34	46.6	219	2	S32078	RNA-directed DNA p	375	34	46.6	342	1	HLHUC4	MHC class I histoc
303	34	46.6	219	2	S32066	RNA-directed DNA p	376	34	46.6	347	2	AC3299	pseudouridylate sy
304	34	46.6	219	2	S32126	RNA-directed DNA p	377	34	46.6	350	2	D71273	probable glycerald
305	34	46.6	219	2	S32074	RNA-directed DNA p	378	34	46.6	350	2	C87536	transcription regu
306	34	46.6	219	2	S32053	RNA-directed DNA p	379	34	46.6	350	2	AG0879	A/G-specific adeni
307	34	46.6	219	2	S32088	RNA-directed DNA p	380	34	46.6	350	2	A40647	methylation-indepe
308	34	46.6	219	2	S32072	RNA-directed DNA p	381	34	46.6	354	2	C83577	hypothetical prote
309	34	46.6	219	2	S32070	RNA-directed DNA p	382	34	46.6	355	2	I80169	class I histocompa
310	34	46.6	219	2	S32087	RNA-directed DNA p	383	34	46.6	361	2	A48445	glyceraldehyde-3-p
311	34	46.6	219	2	S32076	RNA-directed DNA p	384	34	46.6	361	2	T29571	hypothetical prote
312	34	46.6	219	2	S32092	RNA-directed DNA p	385	34	46.6	363	2	JH0542	class I histocompa
313	34	46.6	219	2	S32160	RNA-directed DNA p	386	34	46.6	363	2	S07113	class I histocompa
314	34	46.6	219	2	S32054	RNA-directed DNA p	387	34	46.6	365	2	D95881	probable trehalose
315	34	46.6	219	2	S32135	RNA-directed DNA p	388	34	46.6	366	2	S42823	MHC class I histoc
316	34	46.6	219	2	S32159	RNA-directed DNA p	389	34	46.6	366	2	B37028	MHC class I histoc
317	34	46.6	219	2	S32159	RNA-directed DNA p	390	34	46.6	378	2	D83381	hypothetical prote
318	34	46.6	219	2	S32069	RNA-directed DNA p	391	34	46.6	379	2	I48133	ubiquinol-cytochro
319	34	46.6	219	2	S32128	RNA-directed DNA p	392	34	46.6	379	2	I48132	ubiquinol-cytochro
320	34	46.6	219	2	S32085	RNA-directed DNA p	393	34	46.6	379	2	I48180	ubiquinol-cytochro
321	34	46.6	219	2	S32064	RNA-directed DNA p	394	34	46.6	379	2	I48134	ubiquinol-cytochro

395 34 46.6 386 2 B82921 serine/threonine k
396 34 46.6 399 1 B70936 probable serine/th
397 34 46.6 401 2 G87552 conserved hypothet
398 34 46.6 407 2 T36404 probable monooxyge
399 34 46.6 410 2 H86290 hypothetical prote
400 34 46.6 410 2 G90362 hypothetical prote
401 34 46.6 413 2 JC6512 methyltransferase
402 34 46.6 417 2 F97789 ampG protein [limp
403 34 46.6 424 2 E83442 probable MFS trans
404 34 46.6 431 2 T50177 probable peptide m
405 34 46.6 432 2 AF2801 hypothetical prote
406 34 46.6 432 2 H97580 hypothetical prote
407 34 46.6 442 2 B84232 hypothetical prote
408 34 46.6 446 2 AC2812 conserved hypothet
409 34 46.6 446 2 D97590 hypothetical prote
410 34 46.6 463 2 G90657 hypothetical prote
411 34 46.6 463 2 G85508 hypothetical prote
412 34 46.6 465 2 S41644 polyadenylate-bind
413 34 46.6 470 2 AD0888 SufI protein [limp
414 34 46.6 470 2 E91116 suppressor of ftsI
415 34 46.6 470 2 E85961 suppressor of ftsI
416 34 46.6 470 2 G65088 sufi protein precu
417 34 46.6 472 2 C75505 hypothetical prote
418 34 46.6 472 2 AI3367 multidrug resistan
419 34 46.6 499 2 AC2068 cell death suppres
420 34 46.6 500 2 JC7668 dipeptidyl-peptida
421 34 46.6 518 2 T48143 flavonoid 3',5'-hy
422 34 46.6 519 2 C86160 hypothetical prote
423 34 46.6 526 2 F83166 hypothetical prote
424 34 46.6 541 2 AF2483 hypothetical prote
425 34 46.6 559 2 B47175 reverse transcript
426 34 46.6 559 2 A47175 reverse transcript
427 34 46.6 568 2 E90364 hypothetical prote
428 34 46.6 571 2 AI0506 probable sulfatase
429 34 46.6 584 2 C75364 probable long-chai
430 34 46.6 594 2 G83878 L-lactate permease
431 34 46.6 610 2 AB2436 penicillin-binding
432 34 46.6 623 2 T48859 disease resistance
433 34 46.6 623 2 T06674 hypothetical prote
434 34 46.6 629 2 T16767 hypothetical prote
435 34 46.6 629 2 T06675 hypothetical prote
436 34 46.6 652 2 T16582 hypothetical prote
437 34 46.6 687 2 T27421 hypothetical prote
438 34 46.6 758 2 E83884 hypothetical prote
439 34 46.6 802 2 T45642 FtsH metalloprotei
440 34 46.6 868 2 C81200 aconitate hydratase
441 34 46.6 868 2 H81775 gene UL5 protein -
442 34 46.6 882 1 WMBEU5 replication licens
443 34 46.6 886 1 JC5085 pol polyprotein -
444 34 46.6 902 2 T01668 pol polyprotein -
445 34 46.6 908 2 B69435 signal-transducing
446 34 46.6 912 2 S33980 pol polyprotein -
447 34 46.6 924 2 E71476 alanine-tRNA ligase
448 34 46.6 955 2 T48515 hypothetical prote
449 34 46.6 1002 1 GNLJND HIV-1 retropepsin
450 34 46.6 1002 2 S54378 pol polyprotein -
451 34 46.6 1003 1 GNVWLIV HIV-1 retropepsin
452 34 46.6 1003 1 B44001 HIV-1 retropepsin
453 34 46.6 1003 1 GNVWA2 HIV-1 retropepsin
454 34 46.6 1003 2 T09440 HIV-1 retropepsin
455 34 46.6 1008 2 S38003 pol polyprotein -
456 34 46.6 1012 1 GNVWVL translation elonga
457 34 46.6 1015 1 GNVWH3 HIV-1 retropepsin
458 34 46.6 1027 1 GNLJJSI HIV-1 retropepsin
459 34 46.6 1054 1 GNLJGS HIV-1 retropepsin
460 34 46.6 1061 1 GNLJG4 HIV-1 retropepsin
461 34 46.6 1123 2 T28139 PK4 protein kinase
462 34 46.6 1168 2 T30935 reverse transcript
463 34 46.6 1197 2 D82696 hypothetical prote
464 34 46.6 1205 2 B95845 conserved hypothet
465 34 46.6 1220 1 DJBEC3 DNA-directed DNA p
466 34 46.6 1220 2 T42573 DNA-directed DNA p
467 34 46.6 1271 2 T43269 microcystin synth

468 34 46.6 1411 2 S55123
469 34 46.6 1487 2 F82083
470 34 46.6 1620 2 E83261
471 34 46.6 1707 2 A33526
472 34 46.6 1774 2 S13178
473 34 46.6 2359 2 B96832
474 34 46.6 3036 2 T18995
475 34 46.6 3433 1 GNWVKV genome polyprotein
476 33.5 45.9 191 2 F70694 hypothetical prote
477 33.5 45.9 209 2 B87627 hypothetical prote
478 33.5 45.9 265 2 E87511 conserved hypothet
479 33.5 45.9 344 2 T05437 hypothetical prote
480 33.5 45.9 348 2 H85256 hypothetical prote
481 33.5 45.9 355 2 S62692 L-amino acid oxida
482 33.5 45.9 412 2 AF0668 probable benzoate
483 33.5 45.9 471 2 C82825 UDP-N-acetylmurama
484 33.5 45.9 480 2 T36822 probable transcrip
485 33.5 45.9 519 2 D82536 conserved hypothet
486 33.5 45.9 584 2 D84264 hypothetical prote
487 33.5 45.9 685 2 JC6331 rho-type guanine e
488 33.5 45.9 686 2 E87490 NADH dehydrogenase
489 33 45.2 39 2 I37555 homeobox - human (
490 33 45.2 62 2 C89895 ser protein - huma
491 33 45.2 75 2 S03065 conserved hypothet
492 33 45.2 85 2 AH1010 conserved hypothet
493 33 45.2 94 2 T48710 hypothetical prote
494 33 45.2 114 2 E71171 hypothetical prote
495 33 45.2 142 2 AF0961 heat shock protein
496 33 45.2 150 2 G95348 nitric-oxide reduc
497 33 45.2 154 2 AC0496 heat shock protein
498 33 45.2 159 2 S32327 coat protein - Ses
499 33 45.2 169 2 S26011 hypothetical prote
500 33 45.2 170 2 E83871 hypothetical prote
501 33 45.2 190 2 A82343 conserved hypothet
502 33 45.2 193 2 T35847 probable carbonic
503 33 45.2 195 2 AI2549 molybdopterin bios
504 33 45.2 196 2 AI0502 urease accessory p
505 33 45.2 206 2 AD1898 hypothetical prote
506 33 45.2 207 2 C36961 conserved hypothet
507 33 45.2 208 2 AD2808 methylated-DNA-[pr
508 33 45.2 209 1 XURTMC alpha-ribazole-5'-
509 33 45.2 209 2 A83849 uracil-DNA glycosy
510 33 45.2 213 2 AC3389 hypothetical prote
511 33 45.2 213 2 AH2253 hypothetical prote
512 33 45.2 214 1 ASLJSM vif protein - simi
513 33 45.2 214 1 ASLJSM vif protein - simi
514 33 45.2 217 2 B97587 hypothetical prote
515 33 45.2 218 2 JC7220 nuclear protein SR
516 33 45.2 219 2 S32095 RNA-directed DNA p
517 33 45.2 219 2 A46097 GPI-anchor biosynt
518 33 45.2 224 2 F70614 probable malonyl c
519 33 45.2 229 2 JC7219 nuclear protein SR
520 33 45.2 253 2 JQ2255 tricose-phosphate i
521 33 45.2 257 2 T51696 3-methyl-2-oxobuta
522 33 45.2 260 2 E81194 biotin synthesis p
523 33 45.2 260 2 B83833 hypothetical prote
524 33 45.2 262 2 I40221 divIB protein - Ba
525 33 45.2 264 2 AG2095 hypothetical prote
526 33 45.2 264 2 AH2202 hypothetical prote
527 33 45.2 269 2 S73999 hypothetical prote
528 33 45.2 274 2 H72521 probable thiazole
529 33 45.2 288 2 S39888 SMR2 protein - Pod
530 33 45.2 293 2 AF0617 probable DNA methy
531 33 45.2 295 2 E75366 glutamyl-tRNA synt
532 33 45.2 300 2 S69185 glyceraldehyde-3-p
533 33 45.2 302 2 S75227 hypothetical prote
534 33 45.2 306 2 AE2885 transcription regu
535 33 45.2 306 2 B97661 probable transcrip
536 33 45.2 308 2 B87059 probable pseudouri
537 33 45.2 308 2 A70761 hypothetical prote
538 33 45.2 308 2 S72886 hypothetical prote
539 33 45.2 313 2 T35826 probable dehydroge
540 33 45.2 315 2 E84937 cysteine synthase

541	33	45.2	317	2	G83544	probable transcrip	614	33	45.2	513	1	S50216	translation initia
542	33	45.2	321	2	T42591	gene 48 protein -	615	33	45.2	529	2	E87259	hypothetical prote
543	33	45.2	330	2	F64905	probable sugar tra	616	33	45.2	533	2	S43526	amidophosphoribosy
544	33	45.2	330	2	G85723	probable transport	617	33	45.2	535	2	S76564	hypothetical prote
545	33	45.2	330	2	B90894	probable transport	618	33	45.2	536	2	AH3087	ribitol kinase [im
546	33	45.2	331	1	DEUTC	glyceraldehyde-3-p	619	33	45.2	536	2	A99199	D-ribulokinase (EC
547	33	45.2	332	1	DEJNGI	glyceraldehyde-3-p	620	33	45.2	540	2	S41942	cellulose 1,4-beta
548	33	45.2	334	2	I39602	glyceraldehyde-3-p	621	33	45.2	540	2	S76869	hypothetical prote
549	33	45.2	334	2	G71504	glyceraldehyde-3-p	622	33	45.2	541	2	F84187	phytoene dehydroge
550	33	45.2	335	2	JC7529	glyceraldehyde-3-p	623	33	45.2	542	1	A54963	transcription fact
551	33	45.2	335	2	B72053	glyceraldehyde 3-p	624	33	45.2	551	2	D69282	glutamyl-tRNA synt
552	33	45.2	335	2	E86568	glyceraldehyde-3-p	625	33	45.2	560	2	S41808	glucose-6-phosphat
553	33	45.2	335	2	H81662	glyceraldehyde 3-p	626	33	45.2	567	2	T16105	hypothetical prote
554	33	45.2	335	2	S59579	glyceraldehyde 3-p	627	33	45.2	568	2	T05218	hypothetical prote
555	33	45.2	335	2	T40292	glyceraldehyde-3-p	628	33	45.2	582	2	T16104	hypothetical prote
556	33	45.2	335	2	H70545	glyceraldehyde 3-p	629	33	45.2	594	2	B82994	hypothetical prote
557	33	45.2	336	1	D69231	probable fabH prot	630	33	45.2	600	2	F75424	probable cell cycl
558	33	45.2	336	2	T40235	probable X-Pro dip	631	33	45.2	610	2	S41315	hypothetical prote
559	33	45.2	336	2	H70693	hypothetical prote	632	33	45.2	618	2	AI0171	probable exported
560	33	45.2	336	2	E72389	hypothetical prote	633	33	45.2	625	2	T40742	hypothetical integ
561	33	45.2	337	2	S26974	glyceraldehyde-3-p	634	33	45.2	633	2	B70946	NADH2 dehydrogenas
562	33	45.2	337	2	S26973	glyceraldehyde-3-p	635	33	45.2	634	2	JC4248	calcium binding PW
563	33	45.2	337	2	T27635	homeobox protein c	636	33	45.2	648	1	H69878	probable protein k
564	33	45.2	338	2	AH3321	transcription regu	637	33	45.2	696	2	A91247	phage transposase
565	33	45.2	343	2	AG2126	glyceraldehyde-3-p	638	33	45.2	698	2	T39050	hypothetical prote
566	33	45.2	343	2	AI0581	DNA polymerase III	639	33	45.2	739	2	T45429	polyphosphate kina
567	33	45.2	343	2	B85564	DNA polymerase III	640	33	45.2	742	2	E70673	probable ppk prote
568	33	45.2	343	2	F90713	DNA polymerase III	641	33	45.2	772	2	E82592	transcription-rela
569	33	45.2	343	2	A45251	DNA-directed DNA p	642	33	45.2	775	2	AB3402	topoisomerase IV c
570	33	45.2	347	1	S01340	D-amino-acid oxida	643	33	45.2	807	2	H75634	myosin-Ic - mouse
571	33	45.2	355	2	AG1877	hypothetical prote	644	33	45.2	822	2	E75523	ATP-dependent heli
572	33	45.2	364	2	C84187	hypothetical prote	645	33	45.2	838	2	B83150	probable ATP-depen
573	33	45.2	366	2	T42972	hypothetical prote	646	33	45.2	861	2	T36381	probable large ATP
574	33	45.2	380	2	A71390	ubiquinol-cytochro	647	33	45.2	862	2	T36380	probable large ATP
575	33	45.2	383	2	T20572	hypothetical prote	648	33	45.2	875	2	H81739	alanyl-tRNA synthe
576	33	45.2	384	2	S51796	vasodilator-stimul	649	33	45.2	885	2	G91080	hypothetical prote
577	33	45.2	386	2	A83025	probable acyl-CoA	650	33	45.2	899	2	H85925	hypothetical prote
578	33	45.2	389	2	B69096	corrinoid/iron-sul	651	33	45.2	913	2	T35718	hypothetical prote
579	33	45.2	389	2	S73910	probable serine/th	652	33	45.2	1000	2	T30280	hypothetical prote
580	33	45.2	393	2	JC5614	RNB6 protein - rat	653	33	45.2	1009	2	S44621	C50C3.2 protein -
581	33	45.2	399	2	S75030	hypothetical prote	654	33	45.2	1010	2	T36383	probable large ATP
582	33	45.2	401	2	AF2418	hypothetical prote	655	33	45.2	1012	1	DJBE6S	DNA-directed DNA p
583	33	45.2	408	2	AB6652	LPS biosynthesis p	656	33	45.2	1012	2	T44185	probable DNA-direc
584	33	45.2	408	2	E70127	flagellar motor sw	657	33	45.2	1012	2	T43998	DNA polymerase [im
585	33	45.2	408	2	T50876	hypothetical membr	658	33	45.2	1028	2	A59253	myosin I beta - hu
586	33	45.2	409	2	T51126	hypothetical prote	659	33	45.2	1028	2	S41749	myosin heavy chain
587	33	45.2	411	2	T22290	hypothetical prote	660	33	45.2	1028	2	S37146	myosin I heavy cha
588	33	45.2	414	2	S75052	hypothetical prote	661	33	45.2	1073	2	I51055	recombination acti
589	33	45.2	416	2	E69374	N-ethylammeline ch	662	33	45.2	1078	2	E75407	isoleucyl-tRNA syn
590	33	45.2	416	2	E87286	3-deoxy-D-manno-oc	663	33	45.2	1094	2	F70697	probable arabinosy
591	33	45.2	420	2	E75378	probable valine-py	664	33	45.2	1121	2	G64103	exodeoxyribonuclea
592	33	45.2	420	2	C83246	probable binding p	665	33	45.2	1124	1	GNLJFP	HIV-1 retropepsin
593	33	45.2	430	2	AG0531	cell cycle protein	666	33	45.2	1124	2	B45557	HIV-1 retropepsin
594	33	45.2	431	2	A81150	histidyl-tRNA synt	667	33	45.2	1124	2	S23820	pol polyprotein -
595	33	45.2	432	2	A83060	hypothetical prote	668	33	45.2	1127	2	T21635	hypothetical prote
596	33	45.2	435	2	B72418	conserved hypotHet	669	33	45.2	1137	1	WNBBB1	ribonucleoside-dip
597	33	45.2	440	2	I49681	glyceraldehyde-3-p	670	33	45.2	1143	2	A69465	DNA-directed DNA p
598	33	45.2	448	2	B45438	myosin I beta, MMI	671	33	45.2	1165	2	T21636	hypothetical prote
599	33	45.2	451	2	S77599	probable coproporp	672	33	45.2	1256	2	S14556	asparagine-rich pr
600	33	45.2	451	2	AF3053	conserved hypotHet	673	33	45.2	1319	1	C43735	bcsC protein - Ace
601	33	45.2	451	2	G98232	hypothetical prote	674	33	45.2	1339	2	T47841	hypothetical prote
602	33	45.2	458	2	B75413	major facilitator	675	33	45.2	1404	1	A48196	protein-tyrosine k
603	33	45.2	467	2	A57627	p55 erythrocyte me	676	33	45.2	1451	2	D64203	DNA polymerase III
604	33	45.2	469	2	B70607	probable cyss prot	677	33	45.2	1616	2	T17884	S-layer protein -
605	33	45.2	473	2	C86949	probable cysteinyl	678	33	45.2	1733	2	D70887	probable polyketid
606	33	45.2	480	2	T24087	hypothetical prote	679	33	45.2	1784	2	E86921	polyketide synthas
607	33	45.2	485	2	E83620	probable amidase p	680	33	45.2	1882	2	S73484	hypothetical prote
608	33	45.2	487	2	AI2692	conserved hypotHet	681	33	45.2	2399	2	H71879	toxin-like outer m
609	33	45.2	487	2	E97474	BH3939 hypotHetica	682	33	45.2	3430	1	GNWVWV	genome polyprotein
610	33	45.2	499	2	H83254	probable MFS trans	683	33	45.2	7576	2	T17428	FK506 polyketide s
611	33	45.2	505	2	S68518	tub protein, brain	684	33	45.2	202	2	S56267	probable membrane
612	33	45.2	510	2	S41943	cellulose 1,4-beta	685	32.5	44.5	207	2	T03526	blub protein - Rho
613	33	45.2	511	2	S44716	cellulose 1,4-beta	686	32.5	44.5	219	2	F87444	dethiobiotin synth

687 32.5 44.5 245 2 A84279 hypothetical prote
688 32.5 44.5 417 2 AC0546 probable metabolit
689 32.5 44.5 678 2 D82415 exoribonuclease II
690 32.5 44.5 1736 2 T00391 hypothetical prote
691 32 43.8 62 2 S01004 trypsin inhibitor
692 32 43.8 62 2 A95051 ribosomal protein
693 32 43.8 62 2 F97921 50S ribosomal prot
694 32 43.8 84 2 D83528 hypothetical prote
695 32 43.8 84 2 AB2301 hypothetical prote
696 32 43.8 85 2 B91147 hypothetical prote
697 32 43.8 85 2 F85992 hypothetical prote
698 32 43.8 85 2 C65120 hypothetical 10.0
699 32 43.8 93 2 I47624 Ig heavy chain V-V
700 32 43.8 93 2 E70967 hypothetical prote
701 32 43.8 99 2 AH3429 transposase BMEI14
702 32 43.8 104 2 E82797 conserved hypothet
703 32 43.8 105 2 D72546 hypothetical prote
704 32 43.8 111 2 H95908 conserved hypothet
705 32 43.8 112 2 S75152 hypothetical prote
706 32 43.8 114 2 S22559 Ig heavy chain V r
707 32 43.8 117 2 AB2757 hypothetical prote
708 32 43.8 123 2 T49336 hypothetical prote
709 32 43.8 125 1 A46315 E4 protein - human
710 32 43.8 125 2 F81197 holo-(acyl-carrier
711 32 43.8 132 1 S15618 E4 protein - human
712 32 43.8 133 2 JC2003 NADH ubiquinone ox
713 32 43.8 142 2 B82223 hypothetical prote
714 32 43.8 144 2 G83228 probable transcrip
715 32 43.8 145 2 AG2706 conserved hypothet
716 32 43.8 147 2 T38472 hypothetical prote
717 32 43.8 147 2 A70928 hypothetical prote
718 32 43.8 150 2 AF3634 nitric-oxide reduc
719 32 43.8 153 2 F70950 hypothetical prote
720 32 43.8 158 2 D72305 hypothetical prote
721 32 43.8 167 2 S52220 cobinamide kinase
722 32 43.8 170 1 NWMU2 2S albumin 2 precu
723 32 43.8 177 2 AB2294 hypothetical prote
724 32 43.8 181 2 G70692 hypothetical prote
725 32 43.8 183 2 AC0767 dTDP-4-dehydrorham
726 32 43.8 183 2 S15302 hypothetical prote
727 32 43.8 186 2 G97488 hypothetical prote
728 32 43.8 195 1 SAVLDS delta large antige
729 32 43.8 199 2 S72728 probable L-a-amino
730 32 43.8 200 1 XUSMEG dTDP-4-dehydrorham
731 32 43.8 206 1 NBHUIB platelet glycoprot
732 32 43.8 208 2 JN0638 platelet glycoprot
733 32 43.8 210 2 G83082 hypothetical prote
734 32 43.8 210 2 G85729 hypothetical prote
735 32 43.8 214 2 S07989 vif protein - simi
736 32 43.8 214 2 T11561 vif protein - simi
737 32 43.8 215 2 D70692 hypothetical prote
738 32 43.8 218 2 H83475 probable transcrip
739 32 43.8 219 2 T18541 mofB protein precu
740 32 43.8 220 2 D82139 conserved hypothet
741 32 43.8 221 2 D87469 transcription regu
742 32 43.8 221 2 B36881 MPM2-reactive phos
743 32 43.8 222 2 A81896 hypothetical prote
744 32 43.8 222 2 F81124 conserved hypothet
745 32 43.8 222 2 B90888 hypothetical prote
746 32 43.8 225 2 G83371 probable amino aci
747 32 43.8 228 2 AI2779 nitroreductase [im
748 32 43.8 231 2 D82555 heme ABC transport
749 32 43.8 235 2 T19328 hypothetical prote
750 32 43.8 236 2 H75275 DNA-binding respon
751 32 43.8 238 2 D97559 probable nitroredu
752 32 43.8 244 2 A82316 hypothetical prote
753 32 43.8 249 2 A70840 hypothetical prote
754 32 43.8 252 2 C84522 22 kDa peroxisomal
755 32 43.8 252 2 T31439 probable cobyric a
756 32 43.8 254 2 T35994 probable ABC-type
757 32 43.8 263 2 T35686 phosphatidylglycer
758 32 43.8 267 2 AC3133 hypothetical prote
759 32 43.8 268 2 B42424 chitinase (EC 3.2.

760 32 43.8 269 2 E83240 hypothetical prote
761 32 43.8 272 2 T35231 hypothetical prote
762 32 43.8 278 2 S06416 DNA-directed DNA p
763 32 43.8 278 2 F84127 hypothetical prote
764 32 43.8 281 2 H86159 hypothetical prote
765 32 43.8 284 2 AH0146 probable sulfura
766 32 43.8 286 2 H87292 enoyl-CoA hydratase
767 32 43.8 288 2 S73016 H+-transporting tw
768 32 43.8 291 2 T10966 6-phosphogluconate
769 32 43.8 293 2 T35157 6-phosphogluconate
770 32 43.8 294 2 I40701 glyceraldehyde-3-p
771 32 43.8 294 2 I41220 glyceraldehyde-3-p
772 32 43.8 294 2 I41118 glyceraldehyde-3-p
773 32 43.8 294 2 I41119 glyceraldehyde-3-p
774 32 43.8 294 2 I41222 glyceraldehyde-3-p
775 32 43.8 294 2 I41221 glyceraldehyde-3-p
776 32 43.8 294 2 I41227 glyceraldehyde-3-p
777 32 43.8 294 2 A96155 hypothetical prote
778 32 43.8 299 2 E87049 hypothetical prote
779 32 43.8 299 2 B83888 transcription regu
780 32 43.8 300 2 S75692 hypothetical prote
781 32 43.8 302 2 A70865 probable transfer
782 32 43.8 302 2 F95387 protein imported
783 32 43.8 305 2 AH0863 regulatory protein
784 32 43.8 305 2 F85932 positive regulator
785 32 43.8 305 2 I41065 glycine cleavage s
786 32 43.8 305 2 D91087 positive regulator
787 32 43.8 308 2 A82391 transcription regu
788 32 43.8 309 2 F83434 translocation prot
789 32 43.8 310 2 JC7853 L-fucose-specific
790 32 43.8 312 2 S76507 hypothetical prote
791 32 43.8 315 2 H83276 probable lipase PA
792 32 43.8 315 2 F91250 probable major hea
793 32 43.8 323 2 G83461 hypothetical prote
794 32 43.8 323 2 E95864 probable ABC trans
795 32 43.8 328 2 H90889 hypothetical prote
796 32 43.8 328 2 G64901 ABC-type transport
797 32 43.8 328 2 A85728 hypothetical prote
798 32 43.8 329 1 DEVKGL glyceraldehyde-3-p
799 32 43.8 330 2 T36944 hypothetical prote
800 32 43.8 331 1 DEECG3 glyceraldehyde-3-p
801 32 43.8 331 2 S57279 glyceraldehyde-3-p
802 32 43.8 331 2 S57281 glyceraldehyde-3-p
803 32 43.8 331 2 AG0711 glyceraldehyde 3-p
804 32 43.8 331 2 S57280 glyceraldehyde-3-p
805 32 43.8 331 2 D85788 glyceraldehyde-3-p
806 32 43.8 331 2 H90939 glyceraldehyde-3-p
807 32 43.8 332 1 DEBYG1 glyceraldehyde-3-p
808 32 43.8 332 1 DEBYG2 glyceraldehyde-3-p
809 32 43.8 332 1 DEBYG3 glyceraldehyde-3-p
810 32 43.8 333 1 DEKZGR glyceraldehyde-3-p
811 32 43.8 333 2 JC6310 glyceraldehyde-3-p
812 32 43.8 334 2 D82803 glyceraldehyde-3-p
813 32 43.8 334 2 AI0262 glyceraldehyde-3-p
814 32 43.8 334 2 G86723 exodeoxyribonuclea
815 32 43.8 335 2 S43339 glyceraldehyde-3-p
816 32 43.8 335 2 T12046 glyceraldehyde-3-p
817 32 43.8 337 2 JC5023 CMP-sialic acid tr
818 32 43.8 338 2 S26976 glyceraldehyde-3-p
819 32 43.8 338 2 JN0452 glyceraldehyde-3-p
820 32 43.8 338 2 A44132 D-aspartate oxidas
821 32 43.8 338 2 C83570 hypothetical prote
822 32 43.8 340 2 T09663 glyceraldehyde-3-p
823 32 43.8 340 2 AD0701 tetrathionate redu
824 32 43.8 341 2 T35426 probable oxidoredu
825 32 43.8 347 2 E64069 yceG protein homol
826 32 43.8 350 2 F75259 glutamyl-tRNA redu
827 32 43.8 353 2 B72738 hypothetical prote
828 32 43.8 355 2 C83354 hypothetical prote
829 32 43.8 359 2 D83103 probable phospholi
830 32 43.8 363 2 T45920 hypothetical prote
831 32 43.8 365 2 B83829 glycine oxidase (s
832 32 43.8 368 2 G65119 hypothetical 40.4

833	32	43.8	371	2	A88520	41.8K hypothetical	906	32	43.8	575	2	A11996	hypothetical prote
834	32	43.8	373	2	E70338	probable aspartate	907	32	43.8	583	2	A13098	hypothetical prote
835	32	43.8	376	2	C87596	glycosyl hydrolase	908	32	43.8	591	2	G90988	probable chaperoni
836	32	43.8	377	2	A87288	glutamate 5-kinase	909	32	43.8	592	2	T42078	serine/threonine p
837	32	43.8	377	2	D64385	aspartate transami	910	32	43.8	597	2	JC5829	alkylglycerone-pho
838	32	43.8	380	2	S70964	pkn5 protein - Myx	911	32	43.8	601	2	T36323	probable membrane
839	32	43.8	386	2	H87480	conserved hypothet	912	32	43.8	606	2	G64659	flagellar hook-ass
840	32	43.8	387	2	T01210	glucose-6-phosphat	913	32	43.8	606	2	H98187	prsd protein (U891
841	32	43.8	388	2	G70729	hypothetical prote	914	32	43.8	610	2	T16761	hypothetical prote
842	32	43.8	393	2	B85992	probable transport	915	32	43.8	615	2	A10615	probable exported
843	32	43.8	394	2	C82439	peptide methionine	916	32	43.8	615	2	H90754	probable amidase [
844	32	43.8	396	1	XNECD	aspartate transami	917	32	43.8	615	2	F85618	probable amidase y
845	32	43.8	396	2	AD0616	aspartate aminotra	918	32	43.8	615	2	D64832	ycbB protein precu
846	32	43.8	396	2	A85619	aspartate aminotra	919	32	43.8	619	2	A82638	DNA topoisomerase
847	32	43.8	396	2	C90755	aspartate aminotra	920	32	43.8	622	2	JC7973	synleurin - human
848	32	43.8	396	2	T35254	conserved hypothet	921	32	43.8	634	2	T00054	hypothetical prote
849	32	43.8	399	2	B82723	carbamoyl-phosphat	922	32	43.8	644	1	W1WL58	probable chaperoni
850	32	43.8	399	2	F91146	probable transport	923	32	43.8	646	2	A85834	hypothetical prote
851	32	43.8	403	2	AD0748	tyrosine-specific	924	32	43.8	648	2	E64973	polo-like kinase-1
852	32	43.8	409	2	T47118	thiamine pyridinyl	925	32	43.8	648	2	T43337	hypothetical prote
853	32	43.8	411	1	I55604	platelet glycoprot	926	32	43.8	651	2	T25953	hypothetical prote
854	32	43.8	414	2	C75461	hypothetical prote	927	32	43.8	659	2	E86313	hypothetical prote
855	32	43.8	414	2	S53075	probable membrane	928	32	43.8	660	1	S54746	cytochrome c-type
856	32	43.8	415	2	T38324	probable trna meth	929	32	43.8	664	2	PC4002	phosphatidylinosit
857	32	43.8	416	1	JC4952	transcription init	930	32	43.8	681	1	H82059	2',3'-cyclic-nucle
858	32	43.8	418	2	S46315	aspartate transami	931	32	43.8	682	2	T41680	hypothetical prote
859	32	43.8	424	2	A35861	interferon consens	932	32	43.8	684	2	F70810	hypothetical prote
860	32	43.8	424	2	S44506	regulator protein	933	32	43.8	686	2	T08919	hypothetical prote
861	32	43.8	425	2	A45064	interferon consens	934	32	43.8	691	2	T44543	probable bacteriop
862	32	43.8	427	2	B83001	probable C4-dicarb	935	32	43.8	709	2	F75584	hypothetical prote
863	32	43.8	428	2	B71403	hypothetical prote	936	32	43.8	715	2	S38051	DOA1 protein - yea
864	32	43.8	449	2	D96027	probable nitrilotr	937	32	43.8	719	2	A83800	penicillin-binding
865	32	43.8	452	2	A71877	proline/betaine tr	938	32	43.8	722	2	G86746	hypothetical prote
866	32	43.8	454	2	C82682	glutamate-cysteine	939	32	43.8	725	2	A11544	conserved hypothet
867	32	43.8	456	2	T06136	aspartate transami	940	32	43.8	727	2	T08920	hypothetical prote
868	32	43.8	461	2	H64636	proline/betaine tr	941	32	43.8	728	2	D86278	hypothetical prote
869	32	43.8	469	2	A99656	hypothetical prote	942	32	43.8	735	2	S18623	hypothetical prote
870	32	43.8	473	2	T06167	beta-fructofuranos	943	32	43.8	738	2	C95936	conserved hypothet
871	32	43.8	475	2	T01352	hypothetical prote	944	32	43.8	755	2	D75598	photoreceptor - De
872	32	43.8	482	2	D75346	glutamyl-tRNA(Gln)	945	32	43.8	761	2	T09052	hypothetical prote
873	32	43.8	489	2	A11276	multidrug-efflux t	946	32	43.8	763	2	E96571	hypothetical prote
874	32	43.8	489	2	A11639	multidrug-efflux t	947	32	43.8	765	2	T09574	transport protein
875	32	43.8	493	2	AB0451	conserved hypothet	948	32	43.8	775	2	B64319	carbon-monoxide de
876	32	43.8	493	2	AD1398	drug-export protei	949	32	43.8	783	2	T30644	hypothetical prote
877	32	43.8	493	2	AG1773	drug-export protei	950	32	43.8	798	2	C98069	primosomal replica
878	32	43.8	498	2	AH3456	glutamate synthase	951	32	43.8	801	1	T52538	1-phosphatidylinos
879	32	43.8	499	2	B85507	hypothetical prote	952	32	43.8	808	2	E64914	dimethylsulfoxide
880	32	43.8	502	2	D86578	S/T protein kinase	953	32	43.8	808	2	E90915	probable oxidoredu
881	32	43.8	502	2	G72045	s/t protein kinase	954	32	43.8	808	2	B85764	probable oxidoredu
882	32	43.8	503	2	E72078	probable sodium-tr	955	32	43.8	830	1	A57060	serine/threonine-s
883	32	43.8	503	2	A86544	NADH (ubiquinone)	956	32	43.8	837	2	A97557	dimethylglycine de
884	32	43.8	504	2	A49467	probable sodium-tr	957	32	43.8	837	2	AD2777	dimethylglycine de
885	32	43.8	504	2	C71535	probable NADH (ubi	958	32	43.8	846	2	F75525	outer membrane pro
886	32	43.8	504	2	T32175	occludin - chicken	959	32	43.8	846	2	F75525	cell surface antig
887	32	43.8	515	2	D90048	hypothetical prote	960	32	43.8	875	1	A57080	nikB protein - Esc
888	32	43.8	525	2	C69794	hypothetical prote	961	32	43.8	899	2	B38529	DNA-directed DNA p
889	32	43.8	525	2	C69794	glutamate synthase	962	32	43.8	911	2	S77659	hypothetical prote
890	32	43.8	531	1	XPYS7A	site-specific DNA-	963	32	43.8	914	2	S18942	excinuclease ABC c
891	32	43.8	532	1	E69821	multidrug resistan	964	32	43.8	916	2	H72372	SNF2/Rad54 helicas
892	32	43.8	535	2	B70580	probable UDP-N-ace	965	32	43.8	916	2	G75417	poly(A) polymerase
893	32	43.8	537	2	AE2454	two-component sens	966	32	43.8	942	2	T34847	probable transcrip
894	32	43.8	538	2	H96008	hypothetical prote	967	32	43.8	951	2	T08987	probable cadmium-t
895	32	43.8	540	2	B45665	adult-specific fl.	968	32	43.8	952	2	S32954	hypothetical prote
896	32	43.8	540	2	B87350	hypothetical prote	969	32	43.8	955	2	S52959	male-specific leth
897	32	43.8	544	2	T38469	conserved hypothet	970	32	43.8	971	2	T24866	hypothetical prote
898	32	43.8	552	1	E57987	cytochrome c-type	971	32	43.8	971	2	T14968	phage lambda-relat
899	32	43.8	552	2	D86101	hypothetical prote	972	32	43.8	978	2	A70848	probable membrane
900	32	43.8	552	2	H91260	cytochrome c-type	973	32	43.8	984	2	C84781	hypothetical prote
901	32	43.8	559	2	AI0135	DNA repair protein	974	32	43.8	992	2	A83324	probable sensor/re
902	32	43.8	560	2	S65148	tRNA adenyllyltrans	975	32	43.8	1003	2	T13856	ksr protein - frui
903	32	43.8	564	2	E70394	ABC transporter (h	976	32	43.8	1014	1	DEBY	oxoglutarate dehyd
904	32	43.8	567	2	S69778	adhesin AP65-1 pre	977	32	43.8	1014	2	JE0333	klotho protein - r
905	32	43.8	573	2	I60247	SEC23 protein homo	978	32	43.8	1016	1	S40838	formate dehydrogen

hypothetical prote
hypothetical prote
probable chaperoni
serine/threonine p
alkylglycerone-pho
probable membrane
flagellar hook-ass
prsd protein (U891
hypothetical prote
probable exported
probable amidase [
probable amidase y
ycbB protein precu
DNA topoisomerase
synleurin - human
hypothetical prote
E1 protein - human
probable chaperoni
hypothetical prote
polo-like kinase-1
hypothetical prote
hypothetical prote
cytochrome c-type
phosphatidylinosit
2',3'-cyclic-nucle
hypothetical prote
hypothetical prote
hypothetical prote
probable bacteriop
hypothetical prote
DOA1 protein - yea
penicillin-binding
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
photoreceptor - De
hypothetical prote
hypothetical prote
transport protein
carbon-monoxide de
hypothetical prote
primosomal replica
1-phosphatidylinos
dimethylsulfoxide
probable oxidoredu
probable oxidoredu
serine/threonine-s
dimethylglycine de
dimethylglycine de
outer membrane pro
cell surface antig
nikB protein - Esc
DNA-directed DNA p
hypothetical prote
excinuclease ABC c
SNF2/Rad54 helicas
poly(A) polymerase
probable transcrip
probable cadmium-t
hypothetical prote
male-specific leth
hypothetical prote
phage lambda-relat
probable membrane
hypothetical prote
probable sensor/re
ksr protein - frui
oxoglutarate dehyd
klotho protein - r
formate dehydrogen

979 32 43.8 1016 2 D91231 formate dehydrogen
980 32 43.8 1019 2 T00117 dve protein - frui
981 32 43.8 1094 2 C70612 probable recB prot
982 32 43.8 1102 2 S55100 hypothetical prote
983 32 43.8 1170 2 S03308 cell surface glyco
984 32 43.8 1175 2 T46124 hypothetical prote
985 32 43.8 1190 2 T38636 tat binding homolo
986 32 43.8 1194 1 DJBE28 DNA-directed DNA p
987 32 43.8 1203 2 H87687 helicase, UvrD/Rep
988 32 43.8 1204 2 T18812 hypothetical prote
989 32 43.8 1235 1 DJBE16 DNA-directed DNA p
990 32 43.8 1235 1 DJBEAN DNA-directed DNA p
991 32 43.8 1235 1 DJBEH7 DNA-directed DNA p
992 32 43.8 1235 1 DJBEK1 DNA-directed DNA p
993 32 43.8 1235 1 DJBEV1 DNA-directed DNA p
994 32 43.8 1240 1 DJBE21 DNA-directed DNA p
995 32 43.8 1261 2 G83162 respiratory nitrat
996 32 43.8 1268 2 S52781 neurocan - mouse
997 32 43.8 1299 2 AH2090 two-component hybr
998 32 43.8 1333 2 S30356 CDC25 protein homo
999 32 43.8 1369 2 T03104 tegument protein h
1000 32 43.8 1462 1 S32437 pol polyprotein -

ALIGNMENTS

RESULT 1
A36925
transcription activator LysR-type Cbbr - Xanthobacter flavus
C;Species: Xanthobacter flavus
C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 09-Jul-2004
C;Accession: A36925; S13578; S35408
R;van den Bergh, E.R.E.; Dijkhuizen, L.; Meijer, W.G.
J. Bacteriol. 175, 6097-6104, 1993
A;Title: Cbbr, a LysR-type transcriptional activator, is required for expression of the
A;Reference number: A36925; MUID:94012468; PMID:8407781
A;Accession: A36925
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <VAN>
A;Cross-references: UNIPROT:P25545; UNIPARC:UPI0000127169; EMBL:Z22705; NID:g297851; PID
R;Meijer, W.G.; Arnberg, A.C.; Enequist, H.G.; Terpstra, P.; Lidstrom, M.E.; Dijkhuizen,
Mol. Gen. Genet. 225, 320-330, 1991
A;Title: Identification and organization of carbon dioxide fixation genes in Xanthobacte
A;Reference number: S13573; MUID:91172133; PMID:1900916
A;Accession: S13578
A;Molecule type: DNA
A;Residues: 1-150 <MEI>
A;Cross-references: UNIPARC:UPI00001788AC; EMBL:X17252
C;Genetics:
A;Gene: cbbr
A;Start codon: GTG
C;Superfamily: transcription activator LysR-type
C;Keywords: DNA binding; transcription regulation

Query Match 67.8%; Score 49.5; DB 2; Length 333;
Best Local Similarity 66.7%; Pred. No. 0.99;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 IEG-PTLRQWLARA 14
:| | | :| | | | |
Db 264 VEGLPVVRQWLAVRA 278

RESULT 2
AG0147
probable membrane protein YP01203 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C;Accession: AG0147
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0147
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <KUR>
A;Cross-references: UNIPROT:Q8ZGS7; UNIPARC:UPI00000DC87B; GB:AL590842; PIDN:CAC90042.1;
C;Genetics:
A;Gene: YP01203
C;Superfamily: hypothetical protein ydeD

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PTLRQWLARA 14
| | | | | | | | | |
Db 66 PTLRQWAAASA 76

RESULT 3
D70601
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacteri
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C;Accession: D70601
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70601
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-306 <COL>
A;Cross-references: UNIPROT:O05576; UNIPARC:UPI00000CCAD8; GB:Z94752; GB:AL123456; NID:9
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: galU
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidyltransferase

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 3.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAAR 13
| | | | | | | | | |
Db 290 GPDLRRLVAR 300

RESULT 4
T23485
hypothetical protein K08F4.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T23485
R;Hembry, C.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19746
A;Accession: T23485
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-200 <WIL>
A;Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN00022;
A;Experimental source: clone K08F4
C;Genetics:
A;Gene: CESP:K08F4.11
A;Map position: 4

A;Introns: 45/1; 76/1; 111/3
C;Superfamily: glutathione transferase

Query Match 60.3%; Score 44; DB 2; Length 200;
Best Local Similarity 61.5%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
||| | | | | | | | |
Db 183 IETPKLKEWLAKR 195

RESULT 5
T37464
probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37464
R;Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.
submitted to the EMBL Data Library, June 1997
A;Description: Paraquat mediates differential gene expression in C. elegans.
A;Reference number: Z21702
A;Accession: T37464
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-207 <TAW>
A;Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AAB65419.
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: GST3
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 60.3%; Score 44; DB 2; Length 207;
Best Local Similarity 61.5%; Pred. No. 5.1;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
||| | | | | | | | |
Db 190 IETPKLKEWLAKR 202

RESULT 6
SYECEB
2,3-dihydroxybenzoate-[carrier protein] ligase (EC 6.2.1.-) entE - Escherichia coli (str
N;Alternate names: 2,3-dihydroxybenzoate-AMP ligase [mismomer]; dihydroxybenzoic acid-ac
C;Species: Escherichia coli
C;Date: 31-Dec-1989 #sequence_revision 21-Nov-1997 #text_change 09-Jul-2004
C;Accession: H64792; A48308; A32047; I41058; S08076
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64792
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-536 <BLAT>
A;Cross-references: UNIPROT:P10378; UNIPARC:UPI0000129FD2; GB:AE000165; NID:9
A;Experimental source: strain K-12, substrain MG1655
R;Staab, J.F.; Elkins, M.F.; Earhart, C.F.
FEMS Microbiol. Lett. 59, 15-19, 1989
A;Title: Nucleotide sequence of the Escherichia coli entE gene.
A;Reference number: A48308; MUID:89290355; PMID:2525505
A;Note: in MedLine 89290355 this citation is erroneously given as volume 50 rather than
A;Accession: A48308
A;Molecule type: DNA
A;Residues: 1-368,'ECRRKSTAAR',379-536 <STA>
A;Cross-references: UNIPARC:UPI000016F103; GB:M27490; EMBL:X15058; NID:g41345; PIDN:CAA3
R;Liu, J.; Duncan, K.; Walsh, C.T.
J. Bacteriol. 171, 791-798, 1989
A;Title: Nucleotide sequence of a cluster of Escherichia coli enterobactin biosynthesis
A;Reference number: A91904; MUID:89123155; PMID:2521622

A;Accession: A32047
A;Molecule type: DNA
A;Residues: 393-536 <LIU>
A;Cross-references: UNIPARC:UPI000016F4EA; GB:M24148; NID:g304949; PIDN:AAA16101.1; PID:“
C;Comment: The enzymatic steps in the condensation of L-serine and 2,3-dihydroxybenzoic
ty is based on its recognized homology with 4-coumarate-CoA ligase and by analogy with t;
C;Comment: The formation of 2,3-dihydroxybenzoyl-AMP has been observed. The rapid reacti.
carrier protein) to release AMP, has also been observed.
C;Genetics:
A;Gene: entE
A;Map position: 14 min
C;Function:
A;Description: catalyzes the formation of 2,3-dihydroxybenzoyl-[carrier protein], AMP an
A;Pathway: enterobactin biosynthesis
A;Note: this is one component of a membrane-bound multienzyme complex that catalyzes the
for transport into the cell
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
C;Keywords: acid-thiol ligase; enterobactin biosynthesis; membrane-associated complex
F;69-526/Domain: acetate-CoA ligase homology <ACL>

Query Match 60.3%; Score 44; DB 1; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:: |||||::||
Db 521 VDKKQLRQWLASRA 534

RESULT 7
E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: UNIPARC:UPI00001656E8; GB:AE005174; NID:g12513487; PIDN:AAG54929.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:: |||||::||
Db 521 VDKKQLRQWLASRA 534

RESULT 8
A99708
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A99708
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99708
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-536 <HAY>
A;Cross-references: UNIPROT:Q8XBV3; UNIPARC:UPI0000129FD1; GB:BA0000007; PIDN:BAB34056.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0633
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
: : |||||:|
Db 521 VDKQLRQWLASRA 534

RESULT 9
G71337
probable dimethyladenosine transferase (ksa) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 31-Dec-2004
C;Accession: G71337
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: G71337
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-285 <COL>
A;Cross-references: UNIPROT:O83357; UNIPARC:UPI000012E0CA; GB:AE001213; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0337
C;Superfamily: dimethyladenosine transferase (rRNA adenosine dimethyltransferase)

Query Match 58.9%; Score 43; DB 2; Length 285;
Best Local Similarity 64.3%; Pred. No. 10;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAARA 14
||| |::|||
Db 98 IEGDVLQQWHAATA 111

RESULT 10
B71325
conserved hypothetical protein TP0421 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: B71325
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: B71325
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-683 <COL>
A;Cross-references: UNIPROT:O83436; UNIPARC:UPI00000C0A71; GB:AE001220; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0421

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAAR 13

Db 89 IEAALHQWGAAR 101
||| | || |||
RESULT 11
E84853
hypothetical protein At2g42400 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84853
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>
A;Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI000017A02D; GB:AE002093; NID:g4567312; PII
C;Genetics:
A;Gene: At2g42400
A;Map position: 2

Query Match 57.5%; Score 42; DB 2; Length 473;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEPTLRQWL 10
: || |::||
Db 343 VEGETIREWL 352

RESULT 12
T11560
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)
C;Species: simian immunodeficiency virus SIVsm
A;Variety: strain E543
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11560
R;Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.; N
J. Virol. 71, 1608-1620, 1997
A;Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficien
A;Reference number: Z17285; MUID:97151152; PMID:8995688
A;Accession: T11560
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1019 <HIR>
A;Cross-references: UNIPROT:P89154; UNIPARC:UPI0000105470; EMBL:U72748; NID:g1695908; PII
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; immunodeficiency

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
||| ||||
Db 184 EGPKLRQW 191

RESULT 13
F91171
probable phosphopantetheinyltransferase [imported] - Escherichia coli (strain O157:H7, s
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F91171
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91171
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <HAY>
A;Cross-references: UNIPROT:Q8X5U4; UNIPARC:UPI000012534F; GB:BA000007; PIDN:BAB37765.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs4342

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARA 14
: || : || || ||
Db 27 QGPRRERWLAGRA 39

RESULT 14
F86017
probable phosphopantetheinyltransferase [imported] - Escherichia coli (strain O157:H7, B
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F86017
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F86017
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <STO>
A;Cross-references: UNIPROT:Q8X5U4; UNIPARC:UPI000012534F; GB:AE005174; NID:g12518155; F
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4867

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARA 14
: || : || || ||
Db 27 QGPRRERWLAGRA 39

RESULT 15
S47694
hypothetical 21.8K protein (ftsY-nika intergenic region) - Escherichia coli (strain K-12
N;Alternate names: hypothetical protein o195
C;Species: Escherichia coli
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47694; F65144
R;Plunkett, G.
submitted to the EMBL Data Library, March 1994
A;Reference number: S47666
A;Accession: S47694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <PLU>
A;Cross-references: UNIPROT:P37623; UNIPARC:UPI0000125350; EMBL:U00039; NID:g466582; PID
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65144
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-195 <BLAT>

A;Cross-references: UNIPARC:UPI0000125350; GB:AE000423; GB:U00096; NID:g1789880; PIDN:AA0
A;Experimental source: strain K-12, substrain MGL655
C;Genetics:
A;Gene: yhhU

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARA 14
: || : || || ||
Db 27 QGPRRERWLAGRA 39

RESULT 16
E87575
ABC transporter, ATP-binding protein CC2634 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87575
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: UNIPROT:Q9A535; UNIPARC:UPI00000C77B8; GB:AE005673; NID:g13424211; P
C;Genetics:
A;Gene: CC2634

Query Match 56.2%; Score 41; DB 2; Length 249;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAAR 13
: ||| : || ||
Db 76 QAPTLAPWLSAR 87

RESULT 17
T45453
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacteri
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 05-Oct-2004
C;Accession: T45453
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z22967
A;Accession: T45453
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-306 <JAM>
A;Cross-references: UNIPROT:Q9Z5G1; UNIPARC:UPI00000D438E; EMBL:AL035500; PIDN:CAB36696.
A;Experimental source: cosmid L373
C;Genetics:
A;Note: galU
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidyltransferase

Query Match 56.2%; Score 41; DB 2; Length 306;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
|| || || ||
Db 290 GPDLRRLVER 300

RESULT 18

C24430
glyceraldehyde-3-phosphate dehydrogenase (NADP) (phosphorylating) (EC 1.2.1.13) C, cytosol
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: C24430
R;Shih, M.C.; Lazar, G.; Goodman, H.M.
Cell 47, 73-80, 1986
A;Title: Evidence in favor of the symbiotic origin of chloroplasts: primary structure and
A;Reference number: A90888; MUID:87002494; PMID:3757034
A;Accession: C24430
A;Molecule type: mRNA
A;Residues: 1-326 <SHI>
A;Cross-references: UNIPROT:P09094; UNIPARC:UPI000012AE7D; GB:M14419; NID:g170240; PIDN:
C;Genetics:
A;Gene: GapC
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: cytosol; NADP; oxidative phosphorylation; oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 326;
Best Local Similarity 35.7%; Pred. No. 26;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:|:::|
Db 179 VDGPSMKDWRGGRA 192

RESULT 19
DEPZG
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: S18484
R;Martin, W.; Gierl, A.; Saedler, H.
Nature 339, 46-48, 1989
A;Title: Molecular evidence for pre-Cretaceous angiosperm origins.
A;Reference number: S17991
A;Accession: S18484
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-336 <MAR>
A;Cross-references: UNIPROT:P26519; UNIPARC:UPI000012AE76; EMBL:X60344; NID:g20548; PIDN:
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1991
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F;4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F;153,180/Active site: Cys, His #status predicted

Query Match 56.2%; Score 41; DB 1; Length 336;
Best Local Similarity 35.7%; Pred. No. 27;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:|:::|
Db 189 VDGPSMKDWRGGRA 202

RESULT 20
A35080
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - common ice plant
C;Species: Mesembryanthemum crystallinum (common ice plant)
C;Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 09-Jul-2004
C;Accession: A35080
R;Ostrem, J.A.; Vernon, D.M.; Bohnert, H.J.
J. Biol. Chem. 265, 3497-3502, 1990
A;Title: Increased expression of a gene coding for NAD:glyceraldehyde-3-phosphate dehydrogenase
allium.
A;Reference number: A35080; MUID:90154012; PMID:2303458
A;Accession: A35080
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-337 <OST>
A;Cross-references: UNIPROT:PI7878; UNIPARC:UPI000012AE73; GB:J05223; NID:g167263; PIDN:

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 337;
Best Local Similarity 35.7%; Pred. No. 27;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:|:::|
Db 190 VDGPSMKDWRGGRA 203

RESULT 21
DEIS3C
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - wh
N;Alternate names: triosephosphate dehydrogenase
C;Species: Sinapis alba (white mustard)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: A24796
R;Martin, W.; Cerff, R.
Eur. J. Biochem. 159, 323-331, 1986
A;Title: Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast
A;Reference number: A24796; MUID:87004643; PMID:3530755
A;Accession: A24796
A;Molecule type: mRNA
A;Residues: 1-338 <MAR>
A;Cross-references: UNIPROT:P04796; UNIPARC:UPI000016DF6D; GB:X04301; NID:g21142; PIDN:CA
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F;2-338/Product: glyceraldehyde-3-phosphate dehydrogenase #status experimental <MAT>
F;7-37/Region: beta-alpha-beta NAD nucleotide-binding fold
F;156,183/Active site: Cys, His #status predicted

Query Match 56.2%; Score 41; DB 1; Length 338;
Best Local Similarity 35.7%; Pred. No. 27;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:|:::|
Db 192 VDGPSMKDWRGGRA 205

RESULT 22
JQ1287
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - Ara
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: JQ1287; JS0614
R;Shih, M.C.; Heinrich, P.; Goodman, H.M.
Gene 104, 133-138, 1991
A;Title: Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosol
A;Reference number: JQ1285; MUID:92009205; PMID:1916285
A;Accession: JQ1287
A;Molecule type: DNA
A;Residues: 1-338 <SHI>
A;Cross-references: UNIPROT:P25858; UNIPARC:UPI000016DB30; GB:M64119; NID:g166709; PIDN:Z
A;Accession: JS0614
A;Molecule type: mRNA
A;Residues: 1-338 <SHI1>
A;Cross-references: UNIPARC:UPI000016DB30; GB:M64116; NID:g166705; PIDN:AAA32794.1; PID:5
A;Experimental source: leaf
C;Genetics:
A;Gene: GapC
A;Map position: 3 0.0cM
A;Introns: 2/1; 12/1; 45/3; 84/2; 117/3; 167/2; 187/1; 267/2
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: cytosol; oxidoreductase

Query Match 56.2%; Score 41; DB 2; Length 338;
Best Local Similarity 35.7%; Pred. No. 27;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14

Db 192 VDGPSPMKDWRGGRA 205
::||:::| ||
RESULT 23
B95325
conserved hypothetical protein Sma0937 [imported] - Sinorhizobium meliloti (strain 1021)
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: B95325
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: B95325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-719 <KUR>
A;Cross-references: UNIPROT:Q92ZH9; UNIPARC:UPI00000CB0F3; GB:AE006469; PIDN:AAK65164.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0937
A;Genome: plasmid

Query Match 56.2%; Score 41; DB 2; Length 719;
Best Local Similarity 46.2%; Pred. No. 58;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAAR 13
::|:|||||:
Db 71 LDDPEVRQWLTA 83

RESULT 24
A97501
topoisomerase iv chain a [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C;Accession: A97501
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-750 <KUR>
A;Cross-references: UNIPROT:Q8UG82; UNIPARC:UPI00000D1A6D; GB:AE007869; PIDN:AAK86962.1;
C;Genetics:
A;Gene: AGR_C_2144
A;Map position: circular chromosome
C;Superfamily: Type II topoisomerase, subunit A; phage T4 DNA topoisomerase (ATP-hydroly

Query Match 56.2%; Score 41; DB 2; Length 750;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 GPTLRQWLAARA 14
|:|||||:
Db 721 GEELREWLADRA 732

RESULT 25
AE2719
topoisomerase IV subunit A parC [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C;Accession: AE2719
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2719
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-750 <KUR>
A;Cross-references: UNIPROT:Q8UG82; UNIPARC:UPI00000D1A6D; GB:AE008688; PIDN:AAL42171.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: parC
A;Map position: circular chromosome
C;Superfamily: Type II topoisomerase, subunit A; phage T4 DNA topoisomerase (ATP-hydroly

Query Match 56.2%; Score 41; DB 2; Length 750;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 GPTLRQWLAARA 14
|:|||||:
Db 721 GEELREWLADRA 732

RESULT 26
S25204
srnX protein - Streptomyces ambofaciens
C;Species: Streptomyces ambofaciens
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C;Accession: S25204; S21599
R;Geistlich, M.; Losick, R.; Turner, J.R.; Rao, R.N.
Mol. Microbiol. 6, 2019-2029, 1992
A;Title: Characterization of a novel regulatory gene governing the expression of a polyk
A;Reference number: S25202; MUID:92374852; PMID:1508047
A;Accession: S25204
A;Molecule type: DNA
A;Residues: 1-239 <GEI>
A;Cross-references: UNIPROT:Q00510; UNIPARC:UPI00000BEAE2; EMBL:X63451; NID:g46699; PIDN
C;Genetics:
A;Gene: srnX
F;39-139/Domain: bioC homology <BIOC>

Query Match 54.8%; Score 40; DB 2; Length 239;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 IEGPTLRQWLAARA 14
:|:|:|:
Db 63 VSGLELSEWMAARA 76

RESULT 27
S27491
hypothetical protein A - Bacillus firmus
C;Species: Bacillus firmus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S27491
R;Quirk, P.G.; Krulwich, T.A.
submitted to the EMBL Data Library, October 1991
A;Reference number: S27490
A;Accession: S27491
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-463 <QUI>
A;Cross-references: UNIPROT:P30267; UNIPARC:UPI000013B6AB; GB:L02548; EMBL:M74194; NID:9

Query Match 54.8%; Score 40; DB 2; Length 463;
Best Local Similarity 61.5%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARA 14
||| ||||| |||
Db 296 EGKTSRQWALERA 308

RESULT 28
A81958
probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: A81958
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-530 <PAR>
A;Cross-references: UNIPROT:Q9JWE3; UNIPARC:UPI00000C497F; GB:AL162753; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0414

Query Match 54.8%; Score 40; DB 2; Length 530;
Best Local Similarity 72.7%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLA 11
||| ||||| |||
Db 190 IEMPVLRPWLA 200

RESULT 29
E81015
ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: E81015
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: E81015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <TET>
A;Cross-references: UNIPROT:Q9JXI9; UNIPARC:UPI00000C4849; GB:AE002552; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2026

Query Match 54.8%; Score 40; DB 2; Length 531;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLA 11
||| ||||| |||
Db 191 IEMPVLRPWLA 201

RESULT 30

S30484
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30484
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30484
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
C;Superfamily: pol polyprotein

Query Match 54.8%; Score 40; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGPTLRQW 9
: || |||||
Db 29 MDGPKLRQW 37

RESULT 31
S30483
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30483
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30483
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111
C;Superfamily: pol polyprotein

Query Match 54.8%; Score 40; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGPTLRQW 9
: || |||||
Db 29 MDGPKLRQW 37

RESULT 32
A39707
erythrocyte membrane band 4.2 protein - human
N;Alternate names: pallidin
N;Contains: erythrocyte membrane band 4.2 protein, long splice form; erythrocyte membrane
C;Species: Homo sapiens (man)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: A39707; A34865; B34865; A34883
R;Korngren, C.; Cohen, C.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 4840-4844, 1991
A;Title: Organization of the gene for human erythrocyte membrane protein 4.2: structural
A;Reference number: A39707; MUID:91271288; PMID:2052563
A;Accession: A39707
A;Molecule type: DNA
A;Residues: 1-721 <KOR1>
A;Cross-references: UNIPROT:P16452; UNIPARC:UPI00001608BD; GB:L06519; NID:g306738; PIDN:;
A;Experimental source: cell type erythrocyte; tissue type peripheral blood; tissue lib h
R;Sung, L.A.; Chien, S.; Chang, L.S.; Lambert, K.; Bliss, S.A.; Bouhassira, E.E.; Nagel,
Proc. Natl. Acad. Sci. U.S.A. 87, 955-959, 1990
A;Title: Molecular cloning of human protein 4.2: a major component of the erythrocyte me
A;Reference number: A34865; MUID:90138995; PMID:1689063
A;Accession: A34865

A;Molecule type: mRNA
A;Residues: 1-364,'KRGLPC',371-379,'H',381-405,'L',407-721 <SUN1>
A;Cross-references: UNIPARC:UPI000016AE3E; GB:M30647; NID:g189433; PIDN:AAA36401.1; PID:
A;Accession: B34865
A;Molecule type: mRNA
A;Residues: 1-3,34-364,'KRGLPC',371-379,'H',381-405,'L',407-721 <SUN2>
A;Cross-references: UNIPARC:UPI000016AE3F; GB:M30646; NID:g189435; PIDN:AAA36402.1; PID:
A;Experimental source: isolate Sickle cell patient; cell type reticulocyte
A;Note: parts of this sequence were determined by protein sequencing
R;Korsgren, C.; Lawler, J.; Lambert, S.; Speicher, D.; Cohen, C.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 613-617, 1990
A;Title: Complete amino acid sequence and homologies of human erythrocyte membrane prote
A;Reference number: A34883; MUID:90138879; PMID:2300550
A;Accession: A34883
A;Molecule type: mRNA
A;Residues: 1-3,34-721 <KOR2>
A;Cross-references: UNIPARC:UPI000016A88B; GB:M29399; NID:g182083; PIDN:AAA35798.1; PID:
C;Comment: This protein is a major constituent of the erythrocyte membrane. It appear
C;Genetics:
A;Gene: GDB:EPB42; PA
A;Cross-references: GDB:127385; OMIM:177070
A;Map position: 15q15-15q15
C;Superfamily: protein-glutamine gamma-glutamyltransferase
C;Keywords: alternative splicing; blocked amino end; glycoprotein; lipoprotein; myristyl
F;2-721/Product: erythrocyte membrane band 4.2 protein, long splice form #status predict
F;2-3,34-721/Product: erythrocyte membrane band 4.2 protein, short splice form #status p
F;298-316/Domain: transmembrane #status predicted <TRM>
F;518-520/Region: cell attachment (R-G-D) motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;103,420,447,529,604,705/Binding site: carbohydrate (Aan) (covalent) #status predicted
F;278/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predic

Query Match 54.8%; Score 40; DB 2; Length 721;
Best Local Similarity 70.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PTLRQLAAR 13
| | | | | | |
Db 280 PILRQLTGR 289

RESULT 33
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; UNIPARC:UPI000000A945F; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190

Query Match 54.8%; Score 40; DB 2; Length 1123;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IEGPTLRQLAA 12
: : | | | | | :
Db 200 VQPPTKRQLWS 211

RESULT 34
CGHU2B

collagen alpha 2(IV) chain precursor - human
N;Alternate names: procollagen alpha 2(IV) chain
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A32024; S00007; S02624; S00246; S17678; S16911; B32117; S16877; S00165; S39
R;Hostikka, S.L.; Tryggvason, K.
J. Biol. Chem. 263, 19488-19493, 1988
A;Title: The complete primary structure of the alpha2 chain of human type IV collagen an
A;Reference number: A32024; MUID:89066769; PMID:3198637
A;Accession: A32024
A;Molecule type: mRNA
A;Residues: 1-1712 <HOS1>
A;Cross-references: UNIPROT:P08572; UNIPARC:UPI0000126D42; EMBL:J04210; EMBL:X05610; GB:
R;Hostikka, S.L.; Kurkinen, M.; Tryggvason, K.
FEBS Lett. 216, 281-286, 1987
A;Title: Nucleotide sequence coding for the human type IV collagen alpha-2 chain cDNA re
ated region.
A;Reference number: S00007; MUID:87219158; PMID:3582677
A;Accession: S00007
A;Molecule type: mRNA
A;Residues: 1254-1398,'V',1400-1712 <HOS2>
A;Cross-references: UNIPARC:UPI0000173BE6; EMBL:J04210; EMBL:X05610; GB:M20753; NID:g295
A;Note: 1399-Ile was also found
R;Hostikka, S.L.; Tryggvason, K.
FEBS Lett. 224, 297-305, 1987
A;Title: Extensive structural differences between genes for the alpha(1) and alpha(2) ch
A;Reference number: S02624; MUID:88083553; PMID:2826228
A;Accession: S02624
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1347-1350;1377-1383;1426-1432;1465-1471;1529-1535;1625-1630 <HOS3>
A;Cross-references: UNIPARC:UPI0000142E66; UNIPARC:UPI0000173BE7; UNIPARC:UPI0000173BE8;
A;Note: complete nucleotide sequence not shown
R;Brazel, D.; Pollner, R.; Oberbaeumer, I.; Kuehn, K.
Eur. J. Biochem. 172, 35-42, 1988
A;Title: Human basement membrane collagen (type IV): the amino acid sequence of the alph
A;Reference number: S00246; MUID:88151998; PMID:3345760
A;Accession: S00246
A;Molecule type: mRNA
A;Residues: 1-682,'G',684-1043 <BRA>
A;Cross-references: UNIPARC:UPI0000173BEC; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PI
R;Oberbaeumer, I.
submitted to the EMBL Data Library, June 1987
A;Reference number: S17678
A;Accession: S17678
A;Molecule type: mRNA
A;Residues: 1-470,'P',472-682,'G',684-1043 <OBE>
A;Cross-references: UNIPARC:UPI000016A709; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PI
R;Poeschl, E.; Pollner, R.; Kuehn, K.
EMBO J. 7, 2687-2695, 1988
A;Title: The genes for the alpha1(IV) and alpha2(IV) chains of human basement membrane
A;Reference number: S02738; MUID:89030632; PMID:2846280
A;Accession: S16911
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <POE>
A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:X12784; GB:M36963; NID:g30072; PIDN:CAA
R;Soininen, R.; Huotari, M.; Hostikka, S.L.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 263, 17217-17220, 1988
A;Title: The structural genes for alpha1 and alpha2 chains of human type IV collagen ar
A;Reference number: A92690; MUID:89034231; PMID:3182844
A;Accession: B32117
A;Molecule type: DNA
A;Residues: 1-33 <SOI1>
A;Cross-references: UNIPARC:UPI000016AGF3; EMBL:J04217; EMBL:J05039; NID:g180759; PIDN:
R;Soininen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 264, 13565-13571, 1989
A;Title: Structural organization of the gene for the alpha-1 chain of human type IV coll
A;Reference number: S16876; MUID:89340433; PMID:2701944
A;Accession: S16877
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <SOI2>

A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; NID:g180759; PIDN:AAA53097.1; PI
A;Note: this sequence was submitted to the EMBL Data Library, October 1988
R;Siebold, B.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 168, 569-575, 1987
A;Title: Construction of a model for the aggregation and cross-linking region (7S domain
is region.
A;Reference number: S00165; MUID:88029476; PMID:3117548
A;Accession: S00165
A;Molecule type: protein
A;Residues: 37-247 <SIE1>
A;Cross-references: UNIPARC:UPI0000173BED
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-Gly
R;Eble, J.A.; Golbik, R.; Mann, K.; Kuehn, K.
EMBO J. 12, 4795-4802, 1993
A;Title: The alpha-1-beta-1 integrin recognition site of the basement membrane collagen
A;Reference number: S39614; MUID:94038963; PMID:8223488
A;Accession: S39615
A;Molecule type: protein
A;Residues: 407-570 <EBL>
A;Cross-references: UNIPARC:UPI0000173BEE
R;MacWright, R.S.; Benson, V.A.; Lovello, K.T.; van der Rest, M.; Fietzek, P.P.
Biochemistry 22, 4940-4948, 1983
A;Title: Isolation and characterization of pepsin-solubilized human basement membrane (b
A;Reference number: S16910; MUID:84053346; PMID:6416291
A;Accession: S16912
A;Molecule type: protein
A;Residues: 490-492,'X',494-496;675-677,'G',679-680,'G',682,684-685,'P' <MAC>
A;Cross-references: UNIPARC:UPI0000173BEF; UNIPARC:UPI0000173BF0
A;Experimental source: placenta
R;Glanville, R.W.; Rauter, A.
Hoppe-Seyler's Z. Physiol. Chem. 362, 943-951, 1981
A;Title: Pepsin fragments of human placental basement-membrane collagens showing interrup
A;Reference number: S16908; MUID:82005835; PMID:6792033
A;Accession: B58517
A;Molecule type: protein
A;Residues: 490-492,'X',494-501,'P',503-507;952-957,'X',959-966,'X',968;984-986,'X',988-
81-1185 <GLA>
A;Cross-references: UNIPARC:UPI0000173BF1; UNIPARC:UPI0000173BF2; UNIPARC:UPI0000173BF3;
R;Killen, P.D.; Francomano, C.A.; Yamada, Y.; Modi, W.S.; O'Brien, S.J.
Hum. Genet. 77, 318-324, 1987
A;Title: Partial structure of the human alpha-2(IV) collagen chain and chromosomal local
A;Reference number: S01450; MUID:88085168; PMID:3692475
A;Accession: S01450
A;Molecule type: mRNA
A;Residues: 1040,'L',1042-1398,'V',1400-1418,'M',1420-1635,'V',1637-1712 <KIL>
A;Cross-references: UNIPARC:UPI0000072E68; EMBL:M24766; NID:g537328; PIDN:AAA52043.1; PI
R;Siebold, B.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 176, 617-624, 1988
A;Title: The arrangement of intra- and intermolecular disulfide bonds in the carboxyterm
A;Reference number: S02550; MUID:89005112; PMID:2844531
A;Accession: S02550
A;Molecule type: protein
A;Residues: 1480-1535;1545-1614;1617-1662,'H',1664-1700,'G',1705-1708;1710-1712 <SIE2>
A;Cross-references: UNIPARC:UPI0000173BF6; UNIPARC:UPI0000173BF7; UNIPARC:UPI0000173BF8;
A;Note: the sequence form Fig. 7 is inconsistent with that shown in Fig. 11 in having 17
R;Myers, J.C.; Howard, P.S.; Jelen, A.M.; Dion, A.S.; Macarak, E.J.
J. Biol. Chem. 262, 9231-9238, 1987
A;Title: Duplication of type IV collagen COOH-terminal repeats and species-specific exp
A;Reference number: A27114; MUID:87250571; PMID:2439508
A;Accession: B27114
A;Molecule type: mRNA
A;Residues: 1486-1574,'I',1576-1712 <MYE>
A;Cross-references: UNIPARC:UPI0000173BFB; EMBL:J02760; NID:g180425; PIDN:AAA58422.1; PI
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL4A2
A;Cross-references: GDB:119792; OMIM:120090
A;Map position: 13q34-13q34
A;Introns: 15/2; 33/3; 1347/1; 1380/1; 1429/1; 1468/1; 1532/1; 1527/3 #status incomplete
A;Note: the alpha 1(IV) and alpha 2(IV) chain genes are encoded on opposite strands with
C;Complex: Type IV collagen is a heterotrimer of two alpha 1(IV) chains (see PIR:CGHU4B)
domains (with disulfide and desmosine cross-links), dimeric associations among trimer ca

rupted helical domain (with disulfide and desmosine cross-links).
C;Function:
A;Description: structural component of basement membrane
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: basement membrane; cell binding; coiled coil; extracellular matrix; glycoprotein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1712/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F;29-57/Domain: amino-terminal nonhelical, NH1 <NH1>
F;58-1485/Region: interrupted helical
F;362-364/Region: cell attachment (R-G-D) motif
F;784-786/Region: cell attachment (R-G-D) motif
F;868-870/Region: cell attachment (R-G-D) motif
F;889-891/Region: cell attachment (R-G-D) motif
F;970-972/Region: cell attachment (R-G-D) motif
F;1069-1071/Region: cell attachment (R-G-D) motif
F;1228-1230/Region: cell attachment (R-G-D) motif
F;1452-1454/Region: cell attachment (R-G-D) motif
F;1486-1712/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1495-1593/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1603-1708/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;42,47,51,53,137,483,485/Disulfide bonds: interchain #status predicted
F;57,87,90,102,165,168,225,239,242/Binding site: carboxylate (Lys) (covalent) #status predicted
F;57/Modified site: 5-hydroxylysine (Lys) #status atypical
F;63,75,96,114,120,123,132,150,159,186,189,198,201,213,216,219,496,499,955,964,1103,1115,
F;87,90,102,165,168,225,239,242/Modified site: 5-hydroxylysine (Lys) #status experimental
F;138/Binding site: carboxylate (Asn) (covalent) #status experimental
F;209/Modified site: 4-hydroxyproline (Pro) #status atypical
F;661-681/Disulfide bonds: #status predicted
F;1275/Binding site: carboxylate (Asn) (covalent) #status predicted
F;1504-1590,1537-1593/Disulfide bonds: (or 1504-1593, 1537-1590) #status experimental
F;1549-1555,1658-1665/Disulfide bonds: #status experimental
F;1612-1705,1646-1708/Disulfide bonds: (or 1612-1708, 1646-1705) #status experimental

Query Match 54.8%; Score 40; DB 1; Length 1712;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
: || |||: ||
Db 8 VAGPALRRWL 17

RESULT 35
A84326
hypothetical protein Vngl740c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84326
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84326
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <STO>
A;Cross-references: UNIPROT:Q9HP97; UNIPARC:UPI000006396C; GB:AE004437; NID:g10581200; P
C;Genetics:
A;Gene: VNG1740C

Query Match 54.1%; Score 39.5; DB 2; Length 325;
Best Local Similarity 64.3%; Pred. No. 46;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 IEGPTLRQWLAA 14
: ||| |||||
Db 14 LEGPA-AAWLAARA 26

RESULT 36

S74539
hypothetical protein slr0740 - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74539
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-131 <KAN>
A;Cross-references: UNIPROT:P72684; UNIPARC:UPI00000C09C4; EMBL:D90899; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: *Synechocystis* hypothetical protein slr0740

Query Match 53.4%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTLRQWLA 11
| | | | |
Db 29 POLRQWLA 36

RESULT 37
S46354
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)
C;Species: simian immunodeficiency virus SIVagm
A;Variety: isolate SABD37
C;Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C;Accession: S46354
R;Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;
EMBO J. 13, 2935-2947, 1994
A;Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A;Reference number: S46335; MUID:94298785; PMID:8026477
A;Accession: S46354
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-217 <JIN>
A;Cross-references: UNIPARC:UPI000010A53B; EMBL:U04018; NID:g466250; PIDN:AAA21512.1; PI
A;Experimental source: isolate SABD37; sabaesus monkey
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
: | | | | |
Db 86 DGPRLRQW 93

RESULT 38
I40327
baf protein - *Bordetella pertussis*
C;Species: *Bordetella pertussis*
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40327; S70669
R;DeShazer, D.; Wood, G.E.; Friedman, R.L.
J. Bacteriol. 177, 3801-3807, 1995
A;Title: Identification of a *Bordetella pertussis* regulatory factor required for transcr
A;Reference number: I40327; MUID:95325323; PMID:7601846
A;Accession: I40327
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-267 <RES>
A;Cross-references: UNIPROT:Q45338; UNIPARC:UPI0000038DE7; EMBL:U12020; NID:g687228; PID:
R;Allen, A.; Maskell, D.
Mol. Microbiol. 19, 37-52, 1996
A;Title: The identification, cloning and mutagenesis of a genetic locus required for lip
A;Reference number: S70669; MUID:96419162; PMID:8821935
A;Accession: S70669
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 239-267 <ALL>
A;Cross-references: UNIPARC:UPI000016E756; EMBL:X90711; NID:g992967; PIDN:CAA62242.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C;Genetics:
A;Gene: baf

Query Match 53.4%; Score 39; DB 2; Length 267;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAR 13
| : | | | | |
Db 195 GAIVRQWLAGR 205

RESULT 39
B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - *Leishmania me*
C;Species: *Leishmania mexicana*
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B48445; S25142
R;Hannaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendoes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992
A;Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d
A;Reference number: A48445; MUID:93063042; PMID:1435864
A;Accession: B48445
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-331 <HAN>
A;Cross-references: UNIPROT:Q01558; UNIPARC:UPI000016BF19; EMBL:X65220; NID:g9552; PIDN
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 53.4%; Score 39; DB 2; Length 331;
Best Local Similarity 38.5%; Pred. No. 57;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
: : | | | : |
Db 185 VDGPSLKDWRRGR 197

RESULT 40
A72514
hypothetical protein APE2086 - *Aeropyrum pernix* (strain K1)
C;Species: *Aeropyrum pernix*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 05-Oct-2004
C;Accession: A72514
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: A72514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-331 <KAN>
A;Cross-references: UNIPROT:Q9YA52; UNIPARC:UPI000005E1E5; DDBJ:AP000063; NID:g5105654;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2086
C;Superfamily: PP-loop ATPase, YdaO type

A;Reference number: JS0327; MUID:90122350; PMID:2611042
A;Accession: JS0328
A;Molecule type: DNA
A;Residues: 1-1035 <HAS>
A;Cross-references: UNIPROT:P18042; UNIPARC:UPI0000174A34
A;Note: this sequence was submitted to JIPID, October 1989
C;Comment: Cleavage sites that yield the mature proteins remain to be determined.
C;Genetics:
A;Gene: pol
A;Start codon: ACA
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase;
F;85-183/Product: retropepsin #status predicted <RTP>
F;109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1035;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
:| | | | |
Db 200 DGPRLRQW 207

RESULT 46
GNLJG2
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ROD)
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: B26262
R;Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M.
Nature 326, 662-669, 1987
A;Title: Genome organization and transactivation of the human immuno-deficiency virus ty
A;Reference number: A26262; MUID:87173056; PMID:3031510
A;Accession: B26262
A;Molecule type: DNA
A;Residues: 1-1036 <GUY>
A;Cross-references: UNIPROT:P04584; UNIPARC:UPI0000131EFA; GB:M15390
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucle
F;86-184/Product: retropepsin #status predicted <RTP>
F;110/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1036;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
:| | | | |
Db 201 DGPKLRQW 208

RESULT 47
GNLJST
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ST)
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: B33943
R;Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;
J. Virol. 64, 890-901, 1990
A;Title: Molecular characterization of an attenuated human immunodeficiency virus type 2
A;Reference number: A33943; MUID:90112662; PMID:2296086
A;Accession: B33943
A;Molecule type: genomic RNA
A;Residues: 1-1055 <KUM>
A;Cross-references: UNIPROT:P20876; UNIPARC:UPI0000131EFC
C;Comment: The pol polyprotein contains reverse transcriptase and endonuclease. However,

C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase; revers
F;105-203/Product: retropepsin #status predicted <RTP>
F;129/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1055;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
:| | | | |
Db 220 DGPKLRQW 227

RESULT 48
S53092
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53092
R;Becker, M.; Zorr, B.; Becker, A.; Habermehl, K.O.
submitted to the EMBL Data Library, March 1995
A;Description: Molecular and phylogenetic characterisation of a Guinea-Bissau-derived hu
A;Reference number: S53091
A;Accession: S53092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1055 <BEC>
A;Cross-references: UNIPROT:O73194; UNIPARC:UPI00001785DA; EMBL:Z48731
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 1055;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
:| | | | |
Db 220 DGPKLRQW 227

RESULT 49
S08436
pol polyprotein - human immunodeficiency virus type 2 D205 (fragment)
C;Species: human immunodeficiency virus type 2 D205, HIV-2 D205
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S08436
R;Dietrich, U.; Adamski, M.; Kreutz, R.; Seipp, A.; Kuehnelt, H.; Ruebsamen-Waigmann, H.
Nature 342, 948-950, 1989
A;Title: A highly divergent HIV-2-related isolate.
A;Reference number: S08434; MUID:90081881; PMID:2594088
A;Accession: S08436
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1058 <DIE>
A;Cross-references: UNIPROT:P15833; UNIPARC:UPI0000131EF6; EMBL:X16109
A;Note: this sequence was submitted to the EMBL Data Library, Aug-1989
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 1058;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAAR 13
:| | :| | | :|
Db 221 DGPKIRQWPLSR 232

Search completed: May 12, 2006, 10:52:21
Job time : 21.0915 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-13
Perfect score: 73
Sequence: 1 IEPTLRQWLAARA 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	68.5	53	Q5YUJ5_NOCFA	Q5YUJ5 nocardia fa
2	50	68.5	302	Q742B3_MYCPA	Q742B3 mycobacteri
3	49.5	67.8	333	CBBR_XANFL	P25545 xanthobacte
4	49	67.1	319	Q9RKM5_STRCO	Q9RKM5 streptomyce
5	48	65.8	607	Q9L8D4_POLCB	Q9L8d4 polyangium
6	48	65.8	815	Q4KS46_9VIRU	Q4ks46 orange-spot
7	48	65.8	941	Q8QUJ6_9VIRU	Q8quj6 infectious
8	47	64.4	296	Q8ZGS7_YERPE	Q8zgs7 versinia pe
9	47	64.4	296	Q66D06_YERPS	Q66d06 versinia ps
10	46	63.0	306	Q7D906_MYCTU	Q7d906 mycobacteri
11	46	63.0	306	O05576_MYCTU	O05576 mycobacteri
12	46	63.0	306	Q7U0W3_MYCBO	Q7u0w3 mycobacteri
13	46	63.0	580	Q89RH2_BRAJA	Q89rh2 bradyrhizob
14	45	61.6	245	O66272_9SPHN	O66272 erythroba
15	45	61.6	249	O82989_9SPHN	O82989 erythroba
16	45	61.6	278	Q9XDV0_9SPHN	Q9xdv0 erythroba
17	45	61.6	421	Q7W1X1_BORPA	Q7w1x1 bordetella
18	45	61.6	421	Q7WQU8_BORBR	Q7wqu8 bordetella
19	45	61.6	756	Q885P2_PSESM	Q885p2 pseudomonas
20	44	60.3	207	GST3_CAEEL	O16116 caenorhabdi
21	44	60.3	536	ENTE_ECOS7	Q8xbv3 escherichia
22	44	60.3	536	ENTE_ECOLI	P10378 escherichia
23	44	60.3	536	Q83M10_SHIFL	Q83m10 shigella fl
24	44	60.3	756	Q4ZQD5_PSESY	Q4zqd5 pseudomonas
25	44	60.3	760	Q7G4N4_ORYSA	Q7g4n4 oryza sativ
26	43	58.9	285	KSGA_TREPA	O83357 treponema p
27	43	58.9	297	Q7UQE4_RHOBA	Q7uge4 rhodopirell
28	43	58.9	306	Q8EJ00_SHEON	Q8ej00 shewanella
29	43	58.9	354	Q8ZYT5_PYRAE	Q8zyt5 pyrobaculum
30	43	58.9	377	Q82PX5_STRAW	Q82px5 streptomyce
31	43	58.9	683	O83436_TREPA	O83436 treponema p

32	43	58.9	728	2	Q617Q3_CABBR	Q617q3 caenorhabdi
33	43	58.9	754	2	Q95Y82_CABEL	Q95y82 caenorhabdi
34	43	58.9	1157	2	Q527F0_MAGGR	Q527f0 magnaporthe
35	42	57.5	91	2	Q8Y0I5_RALSO	Q8y0i5 ralstonia s
36	42	57.5	126	2	Q8N9N4_HUMAN	Q8n9n4 homo sapien
37	42	57.5	127	2	Q4SMZ4_TETNG	Q4smz4 tetraodon n
38	42	57.5	252	2	Q8XPQ9_RALSO	Q8xpg9 ralstonia s
39	42	57.5	313	2	P90433_SIVCZ	P90433 chimpanzee
40	42	57.5	325	2	Q855N9_9CAUD	Q855n9 mycobacteri
41	42	57.5	326	2	P95613_RHIGA	P95613 rhizobium g
42	42	57.5	375	2	Q7FAN4_ORYSA	Q7fan4 oryza sativ
43	42	57.5	375	2	Q7XPP6_ORYSA	Q7xpp6 oryza sativ
44	42	57.5	450	2	Q9SLB9_ARATH	Q9slb9 arabidopsis
45	42	57.5	532	2	Q5GW75_XANOR	Q5gw75 xanthomonas
46	42	57.5	586	2	Q9N6P9_LEIMA	Q9n6p9 leishmania
47	42	57.5	1019	1	POL_SIVS4	P12502 simian immu
48	42	57.5	1019	2	P89154_SIVCZ	P89154 chimpanzee
49	42	57.5	1019	2	Q7ZBR7_SIVCZ	Q7zbr7 chimpanzee
50	42	57.5	1019	2	Q7ZBR5_SIVCZ	Q7zbr5 chimpanzee
51	41.5	56.8	400	2	Q5GYS2_XANOR	Q5gys2 xanthomonas
52	41.5	56.8	410	2	Q4UU69_XANCP	Q4uu69 xanthomonas
53	41.5	56.8	410	2	Q8P9L5_XANCP	Q8p9l5 xanthomonas
54	41.5	56.8	427	2	Q8PLE2_XANAC	Q8ple2 xanthomonas
55	41	56.2	75	2	Q98AJ1_RHILO	Q98aj1 rhizobium l
56	41	56.2	89	2	Q50JW6_9ROSI	Q50jw6 shorea parv
57	41	56.2	89	2	Q50JX7_9ROSI	Q50jx7 shorea lepr
58	41	56.2	89	2	Q50JY2_9ROSI	Q50jy2 shorea lepr
59	41	56.2	89	2	Q50K18_9ROSI	Q50k18 shorea curt
60	41	56.2	89	2	Q50K33_9ROSI	Q50k33 shorea acum
61	41	56.2	89	2	Q50K53_9ROSI	Q50k53 shorea acum
62	41	56.2	89	2	Q50K54_9ROSI	Q50k54 shorea hypo
63	41	56.2	130	2	Q6PBK0_BRARE	Q6pbk0 brachydanio
64	41	56.2	136	2	Q5ZFFJ3_POPTN	Q5zffj3 populus tre
65	41	56.2	136	2	Q5ZFFJ8_POPTN	Q5zffj8 populus tre
66	41	56.2	136	2	Q5ZFFJ9_POPTN	Q5zffj9 populus tre
67	41	56.2	136	2	Q5ZFL0_POPTN	Q5zfl0 populus tre
68	41	56.2	136	2	Q5ZFL1_POPTN	Q5zfl1 populus tre
69	41	56.2	136	2	Q5ZFL2_POPTN	Q5zfl2 populus tre
70	41	56.2	136	2	Q5ZFL4_POPTN	Q5zfl4 populus tre
71	41	56.2	136	2	Q5ZFM1_POPTN	Q5zfm1 populus tre
72	41	56.2	136	2	Q5ZFM7_POPTN	Q5zfm7 populus tre
73	41	56.2	136	2	Q5ZFM9_POPTN	Q5zfm9 populus tre
74	41	56.2	136	2	Q5ZFN3_POPTN	Q5zfn3 populus tre
75	41	56.2	136	2	Q5ZFN5_POPTN	Q5zfn5 populus tre
76	41	56.2	136	2	Q5ZFN7_POPTN	Q5zfn7 populus tre
77	41	56.2	136	2	Q5ZFN8_POPTN	Q5zfn8 populus tre
78	41	56.2	137	2	Q9S7Z0_9ROSI	Q9s7z0 manihot pru
79	41	56.2	137	2	Q9S7Z1_MANES	Q9s7z1 manihot esc
80	41	56.2	137	2	Q9SPU6_MANES	Q9spu6 manihot esc
81	41	56.2	137	2	Q4VJ49_9ASTR	Q4vj49 scaevola ch
82	41	56.2	137	2	Q4VJ48_9ASTR	Q4vj48 scaevola co
83	41	56.2	137	2	Q4VJ44_9ASTR	Q4vj44 scaevola ga
84	41	56.2	137	2	Q4VJ41_9ASTR	Q4vj41 scaevola ga
85	41	56.2	137	2	Q4VJ38_9ASTR	Q4vj38 scaevola ki
86	41	56.2	137	2	Q4VJ35_9ASTR	Q4vj35 scaevola mo
87	41	56.2	137	2	Q4VJ32_9ASTR	Q4vj32 scaevola pr
88	41	56.2	153	2	Q80ZR0_MOUSE	Q80zr0 mus musculu
89	41	56.2	160	2	Q56WJ4_ARATH	Q56wj4 arabidopsis
90	41	56.2	163	2	Q4FAA7_BRANA	Q4faa7 brassica na
91	41	56.2	173	2	Q8GZB8_BRARP	Q8gzb8 brassica ra
92	41	56.2	195	1	ACPT_ECO57	Q8x5u4 escherichia
93	41	56.2	195	1	ACPT_ECOLI	P37623 escherichia
94	41	56.2	195	2	Q8FCN3_ECOL6	Q8fcn3 escherichia
95	41	56.2	195	2	Q6N108_RHOPA	Q6n108 rhodopseudo
96	41	56.2	207	2	Q6PKU2_OROMI	Q6pku2 orobanche m
97	41	56.2	209	2	Q6N1X5_RHOPA	Q6nlx5 rhodopseudo
98	41	56.2	219	2	Q8H6A7_ORYSA	Q8h6a7 oryza sativ
99	41	56.2	223	2	Q7XAP7_HOUCO	Q7xap7 houttuynia
100	41	56.2	224	2	Q9F3Q7_STRCO	Q9f3q7 streptomyce
101	41	56.2	244	2	Q9R7K1_9SPHN	Q9r7k1 erythroba
102	41	56.2	245	2	O82987_9SPHN	O82987 erythroba
103	41	56.2	245	2	O82991_9SPHN	O82991 erythroba
104	41	56.2	249	2	Q9A535_CAUCR	Q9a535 caulobacter

105	41	56.2	266	2	Q81D70_BACCR	Q81d70 bacillus ce
106	41	56.2	268	2	Q7Y0S2_LYCES	Q7y0s2 lycopersico
107	41	56.2	294	2	O04891_LYCES	O04891 lycopersico
108	41	56.2	298	2	Q5YPT6_NOCFA	Q5ypt6 nocardia fa
109	41	56.2	306	2	Q9Z5G1_MYCLE	Q9z5g1 mycobacteri
110	41	56.2	308	1	XERC CORGL	Q8nnz9 corynebacte
111	41	56.2	313	2	Q8VWF0_CAPAN	Q8vwp0 capsicum an
112	41	56.2	326	1	G3PC_TOBAC	P09094 nicotiana t
113	41	56.2	336	1	G3PC_PETCR	P26519 petroselinu
114	41	56.2	336	2	Q89ZV8_BACTN	Q89zv8 bacteroides
115	41	56.2	337	1	G3PC_MESCR	P17878 mesembryant
116	41	56.2	337	1	G3PC_SINAL	P04796 sinapis alb
117	41	56.2	337	2	Q6K5G8_ORYSA	P25858 oryza sativ
118	41	56.2	338	1	G3PC_ARATH	P25858 arabidopsis
119	41	56.2	338	1	G3PC_DIACA	P34921 dianthus ca
120	41	56.2	338	2	Q8LAS0_ARATH	Q8las0 arabidopsis
121	41	56.2	338	2	Q8LK04_SOLTU	Q8lk04 solanum tub
122	41	56.2	338	2	Q9FX54_ARATH	Q9fx54 arabidopsis
123	41	56.2	402	2	Q4NC82_9MICC	Q4nc82 arthrobacte
124	41	56.2	485	2	Q8D5X5_VIBVU	Q8d5x5 vibrio vuln
125	41	56.2	631	2	Q6LP74_PHOPR	Q6lp74 photobacter
126	41	56.2	719	2	Q92ZH9_RHIME	Q92zh9 rhizobium m
127	41	56.2	750	2	Q8UG82_AGR75	Q8ug82 agrobacteri
128	41	56.2	766	2	Q9WWZ0_PSEPU	Q9wwz0 pseudomonas
129	41	56.2	776	2	Q6K683_ORYSA	Q6k683 oryza sativ
130	41	56.2	791	2	Q73SJ6_MYCPA	Q73sj6 mycobacteri
131	41	56.2	818	2	Q7WWY0_ALCEU	Q7wwy0 alcaligenes
132	41	56.2	863	2	Q9ST50_MAIZE	Q9st50 zea mays (m
133	41	56.2	1017	2	Q6VG40_SIVCZ	Q6vg40 chimpanzee
134	41	56.2	1153	2	Q92KF2_RHIME	Q92kf2 rhizobium m
135	40	54.8	114	2	Q730Z9_BACC1	Q730z9 bacillus ce
136	40	54.8	115	2	Q81M37_BACAN	Q81m37 bacillus an
137	40	54.8	115	2	Q98HW9_RHILO	Q98hw9 rhizobium l
138	40	54.8	129	2	Q4I3M8_GIBZE	Q4i3m8 gibberella
139	40	54.8	129	2	Q8DX77_SYNEL	Q8dxx7 synechococc
140	40	54.8	137	2	Q6FE11_ACIAI	Q6fe11 acinetobact
141	40	54.8	154	2	Q54XF9_DICDI	Q54xf9 dictyosteli
142	40	54.8	177	2	Q4UTL1_XANCP	Q4utl1 xanthomonas
143	40	54.8	177	2	Q8PA24_XANCP	Q8pa24 xanthomonas
144	40	54.8	194	2	Q72JS0_THET2	Q72js0 thermus the
145	40	54.8	218	2	Q5RJN8_RAT	Q5rjn8 rattus norv
146	40	54.8	230	2	Q72KL1_9HIV1	Q72kl1 human immun
147	40	54.8	238	2	Q835J7_ENTFA	Q835j7 enterococcu
148	40	54.8	239	2	Q00510_STRAM	Q00510 streptomyce
149	40	54.8	241	2	Q4MTI2_BACCE	Q4mti2 bacillus ce
150	40	54.8	241	2	Q631B5_BACCC	Q631b5 bacillus ce
151	40	54.8	242	2	Q81K59_BACAN	Q81k59 bacillus an
152	40	54.8	251	2	Q6HB50_BACHK	Q6hb50 bacillus th
153	40	54.8	251	2	Q815C8_BACCR	Q815c8 bacillus ce
154	40	54.8	271	2	Q4MVQ5_BACCE	Q4mvq5 bacillus ce
155	40	54.8	272	2	Q72XP5_BACC1	Q72xp5 bacillus ce
156	40	54.8	305	2	Q7NRM4_CHRVO	Q7nrm4 chromobacte
157	40	54.8	308	2	Q4NG45_9MICC	Q4ng45 arthrobacte
158	40	54.8	310	2	Q9D8Z7_MOUSE	Q9d8z7 mus musculu
159	40	54.8	331	2	Q63U03_BURPS	Q63u03 burkholderi
160	40	54.8	346	2	Q89M02_BRAJA	Q89m02 bradyrhizob
161	40	54.8	349	2	Q7SKK8_9HIV2	Q7skk8 human immun
162	40	54.8	355	1	TSN10_HUMAN	Q9hiz9 homo sapien
163	40	54.8	369	2	Q4WGC2_ASPFU	Q4wgc2 aspergillus
164	40	54.8	369	2	Q69LD7_ORYSA	Q69ld7 oryza sativ
165	40	54.8	398	2	Q740W3_MYCPA	Q740w3 mycobacteri
166	40	54.8	410	2	Q4LN26_9BURK	Q4ln26 burkholderi
167	40	54.8	410	2	Q63HZ6_BURPS	Q63hz6 burkholderi
168	40	54.8	410	2	Q629V1_BURMA	Q629v1 burkholderi
169	40	54.8	411	2	Q4KDP2_PSEF5	Q4kdp2 pseudomonas
170	40	54.8	429	2	Q7S2Z7_NEUCR	Q7s2z7 neurospora
171	40	54.8	441	2	Q73S55_MYCPA	Q73s55 mycobacteri
172	40	54.8	463	1	YKAA_BACPF	P30267 bacillus ps
173	40	54.8	472	2	Q9WHZ9_9HIV2	Q9whz9 human immun
174	40	54.8	494	2	Q650T8_ORYSA	Q650t8 oryza sativ
175	40	54.8	530	2	Q5F558_NEIG1	Q5f558 neisseria g
176	40	54.8	530	2	Q9JWE3_NEIMA	Q9jwe3 neisseria m
177	40	54.8	531	2	Q9JXI9_NEIMB	Q9jxi9 neisseria m

178	40	54.8	536	2	Q8FK18_ECOL6	Q8fk18 escherichia
179	40	54.8	598	2	Q6N4Y3_RHOPA	Q6n4y3 rhodopseudo
180	40	54.8	619	2	Q4KKX0_HUMAN	Q4kkx0 homo sapien
181	40	54.8	690	1	EPB42_HUMAN	P16452 homo sapien
182	40	54.8	691	2	Q4VB97_HUMAN	Q4vb97 homo sapien
183	40	54.8	818	2	Q5NV45_9BURK	Q5nv45 ralstonia m
184	40	54.8	992	2	Q4P0J6_USTMA	Q4p0j6 ustilago ma
185	40	54.8	1082	2	Q756B5_ASHGO	Q756b5 ashbya goss
186	40	54.8	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis
187	40	54.8	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis
188	40	54.8	1444	2	Q4PG64_USTMA	Q4pg64 ustilago ma
189	40	54.8	1706	2	Q5VZA9_HUMAN	Q5vza9 homo sapien
190	40	54.8	1712	1	CO4A2_HUMAN	P08572 homo sapien
191	40	54.8	1896	2	Q9DRA1_9VIRU	Q9dra1 botrytis vi
192	39.5	54.1	325	2	Q9HP97_HALSA	Q9hp97 halobacteri
193	39.5	54.1	613	2	Q89HC1_BRAJA	Q89hc1 bradyrhizob
194	39	53.4	65	2	Q984B5_RHILO	Q984b5 rhizobium l
195	39	53.4	131	2	P72684_SINY3	P72684 synechocyst
196	39	53.4	133	2	Q4ZPV7_PSESY	Q4zpv7 pseudomonas
197	39	53.4	134	2	Q886I5_PSESM	Q886i5 pseudomonas
198	39	53.4	148	2	Q4UK00_RICFE	Q4uk00 rickettsia
199	39	53.4	154	2	Q9WI01_9HIV2	Q9wi01 human immun
200	39	53.4	160	2	Q73UZ7_MYCPA	Q73uz7 mycobacteri
201	39	53.4	164	2	Q6MNC9_BDEBA	Q6mnc9 bdellovibri
202	39	53.4	167	2	Q888B5_PSESM	Q888b5 pseudomonas
203	39	53.4	168	2	Q9V492_DROME	Q9v492 drosophila
204	39	53.4	169	2	Q4ZXW1_PSESY	Q4zxw1 pseudomonas
205	39	53.4	196	2	Q8SAP5_9BRYO	Q8sap5 mitthyridiu
206	39	53.4	217	2	Q87115_SIVCZ	Q87115 chimpanzee
207	39	53.4	252	2	Q8XY68_RALSO	Q8xy68 ralstonia s
208	39	53.4	259	2	Q8PFI7_XANAC	Q8pfi7 xanthomonas
209	39	53.4	267	1	BAF_BORBR	Q8pf17 bordetella
210	39	53.4	267	1	BAF_BORPE	Q45338 bordetella
211	39	53.4	287	2	Q74Z22_MYCPA	Q74z22 mycobacteri
212	39	53.4	310	2	Q7SKI9_9HIV2	Q7ski9 human immun
213	39	53.4	319	2	Q89HR9_BRAJA	Q89hr9 bradyrhizob
214	39	53.4	326	2	Q98VC4_9HEPC	Q98vc4 hepatitis c
215	39	53.4	330	1	G3PC_LEIME	Q01558 leishmania
216	39	53.4	331	2	Q9YA52_AERPE	Q9ya52 aeropyrum p
217	39	53.4	335	2	Q82D06_STRAW	Q82d06 streptomyce
218	39	53.4	337	1	G3P_PHACH	Q01982 phanerochae
219	39	53.4	338	2	Q4PLN5_9BURK	Q4pln5 ustilago ma
220	39	53.4	338	2	Q8LPN5_9BURK	Q8lpn5 burkholderi
221	39	53.4	340	2	Q8UN04_SIVCZ	Q8un04 chimpanzee
222	39	53.4	340	2	Q8UN03_SIVCZ	Q8un03 chimpanzee
223	39	53.4	343	2	Q4LXZ8_9BURK	Q4lxz8 burkholderi
224	39	53.4	349	2	Q7SKK9_9HIV2	Q7skk9 human immun
225	39	53.4	349	2	Q7SKK6_9HIV2	Q7skk6 human immun
226	39	53.4	349	2	Q7SKK4_9HIV2	Q7skk4 human immun
227	39	53.4	349	2	Q7SKJ6_9HIV2	Q7skj6 human immun
228	39	53.4	349	2	Q7SKJ5_9HIV2	Q7skj5 human immun
229	39	53.4	349	2	Q7SKJ4_9HIV2	Q7skj4 human immun
230	39	53.4	349	2	Q7SKJ3_9HIV2	Q7skj3 human immun
231	39	53.4	349	2	Q7SKJ2_9HIV2	Q7skj2 human immun
232	39	53.4	349	2	Q7SKJ1_9HIV2	Q7skj1 human immun
233	39	53.4	349	2	Q7SKJ0_9HIV2	Q7skj0 human immun
234	39	53.4	349	2	Q7SKI8_9HIV2	Q7ski8 human immun
235	39	53.4	349	2	Q7SKI7_9HIV2	Q7ski7 human immun
236	39	53.4	349	2	Q7SKI6_9HIV2	Q7ski6 human immun
237	39	53.4	349	2	Q7SKI5_9HIV2	Q7ski5 human immun
238	39	53.4	349	2	Q7SKI4_9HIV2	Q7ski4 human immun
239	39	53.4	349	2	Q7SKI3_9HIV2	Q7ski3 human immun
240	39	53.4	349	2	Q7SKI2_9HIV2	Q7ski2 human immun
241	39	53.4	349	2	Q7SKI1_9HIV2	Q7ski1 human immun
242	39	53.4	351	2	Q9E3T1_9HIV2	Q9e3t1 human immun
243	39	53.4	351	2	Q9E3T2_9HIV2	Q9e3t2 human immun
244	39	53.4	351	2	Q9E3T4_9HIV2	Q9e3t4 human immun
245	39	53.4	353	2	Q5AY13_EMENI	Q5ay13 aspergillus
246	39	53.4	353	2	Q9KWU0_9SPHN	Q9kwu0 sphingomona
247	39	53.4	354	2	Q9KZQ9_STRCO	Q9kzq9 streptomyce
248	39	53.4	374	2	Q73S53_MYCPA	Q73s53 mycobacteri
249	39	53.4	396	2	Q90PUI_SIVCZ	Q90pul chimpanzee
250	39	53.4	398	2	Q820L4_NITEU	Q820l4 nitrosomona

Q8fk18	escherichia
Q6n4y3	rhodopseudo
Q4kkx0	homo sapien
P16452	homo sapien
Q4vb97	homo sapien
Q5nv45	ralstonia m
Q4p0j6	ustilago ma
Q756b5	ashbya goss
Q9se99	arabidopsis
Q9spu7	arabidopsis
Q4pg64	ustilago ma
Q5vza9	homo sapien
P08572	homo sapien
Q9dra1	botrytis vi
Q9hp97	halobacteri
Q89hc1	bradyrhizob
Q984b5	rhizobium l
P72684	synechocyst
Q4zpv7	pseudomonas
Q886i5	pseudomonas
Q4uk00	rickettsia
Q9wi01	human immun
Q73uz7	mycobacteri
Q6mnc9	bdellovibri
Q888b5	pseudomonas
Q9v492	drosophila
Q4zxw1	pseudomonas
Q8sap5	mitthyridiu
Q87115	chimpanzee
Q8xy68	ralstonia s
Q8pf17	xanthomonas
Q45338	bordetella
Q74222	mycobacteri
Q7ski9	human immun
Q89hr9	bradyrhizob
Q98vc4	hepatitis c
Q01558	leishmania
Q9ya52	aeropyrum p
Q82d06	streptomyce
Q01982	phanerochae
Q4pln5	ustilago ma
Q8lpn5	burkholderi
Q8un04	chimpanzee
Q8un03	chimpanzee
Q4lxz8	burkholderi
Q7skk9	human immun
Q7skk6	human immun
Q7skk4	human immun
Q7skj6	human immun
Q7skj5	human immun
Q7skj4	human immun
Q7skj3	human immun
Q7skj2	human immun
Q7skj1	human immun
Q7skj0	human immun
Q7ski8	human immun
Q7ski7	human immun
Q7ski6	human immun
Q7ski5	human immun
Q7ski4	human immun
Q7ski3	human immun
Q7ski2	human immun
Q7ski1	human immun
Q9e3t1	human immun
Q9e3t2	human immun
Q9e3t4	human immun
Q5ay13	aspergillus
Q9kwu0	sphingomona
Q9kzq9	streptomyce
Q73s53	mycobacteri
Q90pul	chimpanzee
Q820l4	nitrosomona

251	39	53.4	400	2	069568_MYCLE	069568_mycobacteri	324	39	53.4	1677	2	Q4SDD0_TETNG	Q4sdd0_tetraodon n
252	39	53.4	402	2	Q82R87_STRAW	Q82r87_streptomyce	325	39	53.4	1722	2	Q53L67_ORYSA	Q53l67_oryza sativ
253	39	53.4	403	2	Q88NU2_PSEPK	Q88nu2_pseudomonas	326	39	53.4	2355	2	Q4KCD0_PSEF5	Q4kcd0_pseudomonas
254	39	53.4	414	2	Q8FT01_COREF	Q8ft01_corynebacte	327	39	53.4	3433	1	PCX_DROME	P18490_drosophila
255	39	53.4	457	2	Q7ULE0_RHOBA	Q7ule0_rhodopirell	328	39	53.4	6481	2	Q4QE92_LEIMA	Q4qe92_leishmania
256	39	53.4	468	2	Q88H13_PSEPK	Q88hl3_pseudomonas	329	38.5	52.7	535	2	Q5MZW6_SYNp6	Q5mzw6_synecchococ
257	39	53.4	476	2	Q9WI00_9HIV2	Q9wi00_human immun	330	38	52.1	61	2	Q5YS14_NOCFA	Q5ys14_nocardia fa
258	39	53.4	496	2	Q6CER7_YARLI	Q6cer7_yarrowia li	331	38	52.1	67	2	Q5ISF8_9HOMO	Q5isf8_leccinum sc
259	39	53.4	545	2	Q6N594_RHOPA	Q6n594_rhodopseudo	332	38	52.1	87	2	Q8XZR4_RALSO	Q8xzz4_ralstonia s
260	39	53.4	553	2	Q7WGH7_BORBR	Q7wgh7_bordetella	333	38	52.1	88	2	Q8Y0T1_RALSO	Q8y0t1_ralstonia s
261	39	53.4	553	2	Q7W512_BORPA	Q7w512_bordetella	334	38	52.1	106	2	Q6H0G4_9HOMO	Q6h0g4_leccinum qu
262	39	53.4	553	2	Q7VVX5_BORPE	Q7w512_bordetella	335	38	52.1	106	2	Q6H0J0_9HOMO	Q6h0j0_leccinum ma
263	39	53.4	553	2	Q06347_9HIV2	Q7vvx5_bordetella	336	38	52.1	118	2	Q6H0G2_9HOMO	Q6h0g2_leccinum ca
264	39	53.4	560	2	Q06347_9HIV2	Q06347_human immun	337	38	52.1	120	2	Q5I5E9_9HOMO	Q5ise9_leccinum va
265	39	53.4	571	2	Q7S508_NEUCR	Q06347_human immun	338	38	52.1	123	2	Q5ISF7_9HOMO	Q5isf7_leccinum sc
266	39	53.4	573	2	Q5FUI6_GLUOX	Q7s508_neurospora	339	38	52.1	124	2	Q6H0H8_9HOMO	Q6h0h8_leccinum ro
267	39	53.4	579	2	Q6WLD7_9BACT	Q5fui6_gluconobact	340	38	52.1	129	2	Q6H0G0_9HOMO	Q6h0g0_leccinum mo
268	39	53.4	582	2	Q8AB47_BACTN	Q6wld7_uncultured	341	38	52.1	129	2	Q6IJ60_DROME	Q6ij60_drosophila
269	39	53.4	586	2	Q828W4_STRAW	Q8ab47_bacterioides	342	38	52.1	132	2	Q6H0F7_9HOMO	Q6h0f7_leccinum po
270	39	53.4	600	2	Q9HYJ8_PSEAE	Q828w4_streptomyce	343	38	52.1	132	2	Q8N1Q4_HUMAN	Q8nlq4_homo sapien
271	39	53.4	609	2	Q856X8_9CAUD	Q856x8_mycobacteri	344	38	52.1	133	2	Q5I5G9_9HOMO	Q5i5g9_leccinum sc
272	39	53.4	609	2	Q989R7_RHILO	Q989r7_rhizobium l	345	38	52.1	133	2	Q944T3_FRAAN	Q944t3_fragaria an
273	39	53.4	612	2	Q830J7_ENTFA	Q830j7_enterococcu	346	38	52.1	133	2	Q944T7_FRAAN	Q944t7_fragaria an
274	39	53.4	632	2	Q4LQV6_9BURK	Q4lqv6_burkholderi	347	38	52.1	134	2	Q9RW17_DEIRA	Q9rw17_deinococcus
275	39	53.4	647	2	Q6FV06_CANGA	Q75ij9_oryza sativ	348	38	52.1	140	2	Q9RW17_DEIRA	Q9rw17_deinococcus
276	39	53.4	693	2	Q5TNY2_ANOGA	Q6fv06_candida gla	349	38	52.1	140	2	Q5ISF6_9HOMO	Q5isf6_leccinum sc
277	39	53.4	706	2	Q63U33_BURPS	Q5tny2_anopheles g	350	38	52.1	140	2	Q5I5H2_9HOMO	Q5i5h2_leccinum ro
278	39	53.4	706	2	Q62KA1_BURMA	Q63u33_burkholderi	351	38	52.1	141	2	Q6H0F8_9HOMO	Q6h0f8_leccinum in
279	39	53.4	713	2	Q7NL87_GLOVI	Q62ka1_burkholderi	352	38	52.1	142	2	Q5ISF4_9HOMO	Q5isf4_leccinum sc
280	39	53.4	725	2	Q8WR00_LEIMA	Q7nl87_gloeobacter	353	38	52.1	142	2	Q5ISF1_9HOMO	Q5isf1_leccinum va
281	39	53.4	733	2	Q5XQF4_MAIZE	Q8wr00_leishmania	354	38	52.1	144	2	Q7XWC2_ORYSA	Q7xwc2_oryza sativ
282	39	53.4	740	2	Q8XVE2_RALSO	Q5xqf4_zea mayas (m	355	38	52.1	146	2	Q6H0G3_9HOMO	Q6h0g3_leccinum au
283	39	53.4	766	2	Q9R9Y9_PSEPU	Q8xve2_ralstonia s	356	38	52.1	149	2	Q9L9U0_9ENTR	Q9l9u0_brenneria q
284	39	53.4	766	2	Q33466_PSEPU	Q9r9y9_pseudomonas	357	38	52.1	194	2	Q81924_CICAR	Q81924_cicer ariet
285	39	53.4	766	2	Q88KB4_PSEPK	Q33466_pseudomonas	358	38	52.1	197	1	BETI_PSEAE	Q88af0_pseudomonas
286	39	53.4	788	2	Q89QX5_BRAJA	Q88kb4_pseudomonas	359	38	52.1	197	1	BETI_PSEEM	Q88ap2_periploca s
287	39	53.4	791	2	Q9KU26_VIBCH	Q89qx5_bradyrhizob	360	38	52.1	197	2	Q68AP3_9GENT	Q68ap3_periploca s
288	39	53.4	925	2	Q7XS10_ORYSA	Q9ku26_vibrio chol	361	38	52.1	197	2	Q4ZM61_PSESY	Q4zm61_pseudomonas
289	39	53.4	986	2	Q57059_SIVCZ	Q7xsi0_oryza sativ	362	38	52.1	197	2	Q4K4K9_PSEF5	Q4k4k9_pseudomonas
290	39	53.4	1005	2	Q6Y8X5_9HIV1	Q57059_chimpanzee	363	38	52.1	214	2	Q6H0H9_9HOMO	Q6h0h9_leccinum ve
291	39	53.4	1015	2	Q8JAH1_SIVCZ	Q6y8x5_human immun	364	38	52.1	218	2	Q5P3S9_AZOSE	Q5p3s9_azocarcus sp
292	39	53.4	1022	1	POL_SIVSP	Q8jah1_chimpanzee	365	38	52.1	225	2	Q6H0H1_9HOMO	Q6h0h1_leccinum pa
293	39	53.4	1022	2	Q90317_SIVCZ	P19505_simian immu	366	38	52.1	227	2	Q6H0H6_9HOMO	Q6h0h6_leccinum va
294	39	53.4	1022	2	Q87956_SIVCZ	Q90317_chimpanzee	367	38	52.1	227	2	Q9RBZ4_PSESG	Q9rbz4_pseudomonas
295	39	53.4	1022	2	Q87965_SIVCZ	Q87956_chimpanzee	368	38	52.1	227	2	Q88BV4_PSESM	Q88bv4_pseudomonas
296	39	53.4	1022	2	Q87965_SIVCZ	Q87965_chimpanzee	369	38	52.1	231	2	Q5I5E7_9HOMO	Q5ise7_leccinum va
297	39	53.4	1022	2	Q89620_SIVCZ	Q89620_chimpanzee	370	38	52.1	231	2	Q5I5H5_9HOMO	Q5ish5_leccinum ro
298	39	53.4	1034	1	Q88135_SIVCZ	Q88135_chimpanzee	371	38	52.1	233	2	Q5I5F5_9HOMO	Q5isf5_leccinum sc
299	39	53.4	1035	2	POL_HV2CA	P24107_human immun	372	38	52.1	233	2	Q5I5G4_9HOMO	Q5isg4_leccinum me
300	39	53.4	1035	2	Q73194_9HIV2	Q73194_human immun	373	38	52.1	233	2	Q6H0G7_9HOMO	Q6h0g7_leccinum sc
301	39	53.4	1035	2	Q6R793_9HIV2	Q6r793_human immun	374	38	52.1	233	2	Q6H0H5_9HOMO	Q6h0h5_leccinum sn
302	39	53.4	1036	1	Q6R784_9HIV2	Q6r784_human immun	375	38	52.1	237	2	Q5I5H4_9HOMO	Q5ish4_leccinum ro
303	39	53.4	1049	1	POL_HV2RO	P04584_human immun	376	38	52.1	238	2	Q6H0J2_9HOMO	Q6h0j2_leccinum mo
304	39	53.4	1049	2	POL_HV2G1	P18042_human immun	377	38	52.1	238	2	O12158_9HIV1	O12158_human immun
305	39	53.4	1055	1	Q9XBP6_MYXXA	Q9xbp6_myxococcus	378	38	52.1	239	2	Q5I5H3_9HOMO	Q5ish3_leccinum ro
306	39	53.4	1055	1	POL_HV2ST	P20876_human immun	379	38	52.1	239	2	Q7W1D9_BORPA	Q7w1d9_bordetella
307	39	53.4	1056	2	Q04097_SIVCZ	Q9ytu0_human immun	380	38	52.1	240	2	Q5I5F9_9HOMO	Q5isf9_leccinum me
308	39	53.4	1058	1	POL_HV2D2	Q04097_chimpanzee	381	38	52.1	241	2	Q5D9T2_SCHJA	Q5d9t2_schistosoma
309	39	53.4	1059	2	Q76630_9HIV2	P15833_human immun	382	38	52.1	242	2	Q5I5G2_9HOMO	Q5isg2_leccinum sc
310	39	53.4	1060	2	Q5QGQ1_SIVCZ	Q76630_human immun	383	38	52.1	243	2	Q5I5G7_9HOMO	Q5isg7_leccinum sc
311	39	53.4	1060	2	Q5QFT5_SIVCZ	Q5ggq1_chimpanzee	384	38	52.1	243	2	Q5I5H0_9HOMO	Q5ish0_leccinum ro
312	39	53.4	1060	2	Q5QFR3_SIVCZ	Q5qft5_chimpanzee	385	38	52.1	244	2	Q5I5G7_9HOMO	Q5isg7_leccinum sc
313	39	53.4	1060	2	Q5QFL2_SIVCZ	Q5qfr3_chimpanzee	386	38	52.1	244	2	Q6G269_9SPHN	Q6g269_erythromicr
314	39	53.4	1060	2	Q5QFJ6_SIVCZ	Q5qfl2_chimpanzee	387	38	52.1	244	2	Q7W0F8_BORPE	Q7w0f8_bordetella
315	39	53.4	1098	2	Q70TF0_PHYPA	Q5qfj6_chimpanzee	388	38	52.1	245	2	O66278_9SPHN	O66278_agrobacteri
316	39	53.4	1209	2	Q4WNF8_ASPFU	Q70tf0_physcomitre	389	38	52.1	245	2	Q9ZN87_9SPHN	Q9zn87_porphyrbac
317	39	53.4	1226	2	Q6H0K6_9HIV2	Q4wnf8_aspergillus	390	38	52.1	246	2	Q6G276_9SPHN	Q6g276_porphyrbac
318	39	53.4	1242	1	NPHN_MOUSE	Q6h0k6_human immun	391	38	52.1	246	2	Q66C35_YERPS	Q66c35_yersinia ps
319	39	53.4	1256	2	Q9JIX1_MOUSE	Q9qzs7_mus musculu	392	38	52.1	246	2	Q8ZFW2_YERPE	Q8zfw2_yersinia pe
320	39	53.4	1414	2	Q7SI31_NEUCR	Q9jix1_mus musculu	393	38	52.1	251	2	Q6H0J4_9HOMO	Q6h0j4_leccinum ul
321	39	53.4	1431	2	Q5B491_EMENI	Q7si31_neurospora	394	38	52.1	256	2	Q6H0G6_9HOMO	Q6h0g6_leccinum sc
322	39	53.4	1435	2	Q4WIA7_ASPFU	Q5b491_aspergillus	395	38	52.1	261	2	Q73WB7_MYCPA	Q73wb7_mycobacteri
323	39	53.4	1447	2	Q4FWA4_LEIMA	Q4wia7_aspergillus	396	38	52.1	262	2	Q6H0I1_9HOMO	Q6h0i1_leccinum ve
						Q4fwa4_leishmania							

543	37.5	51.4	555	2	Q7QHJ7_ANOGA	Q7qhj7 anopheles g	616	37	50.7	233	2	Q9QDG2_9HIV1	Q9qdg2 human immun
544	37	50.7	79	2	Q533Q1_9ROSI	Q533q1 quercus gla	617	37	50.7	234	2	Q4K3V9_PSEF5	Q4k3v9 pseudomonas
545	37	50.7	79	2	Q533R2_9ROSI	Q533r2 quercus gla	618	37	50.7	235	2	Q9M4V4_WHEAT	Q9m4v4 triticum ae
546	37	50.7	81	2	Q9NDL7_HYDMA	Q9ndl7 hydra magni	619	37	50.7	235	2	Q4J4K4_AZOVI	Q4j4k4 azotobacter
547	37	50.7	85	2	Q64GN7_9HIV1	Q64gn7 human immun	620	37	50.7	237	1	PYRF_GLOVI	Q7nk22 gloeobacter
548	37	50.7	87	2	Q65930_CAEV	Q65930 caprine art	621	37	50.7	239	2	Q5MI13_9HIV1	Q5mi13 human immun
549	37	50.7	87	2	Q9WR47_9RETR	Q9wr47 brazilian c	622	37	50.7	239	2	Q5MI07_9HIV1	Q5mi07 human immun
550	37	50.7	103	2	Q95QV6_CAEEL	Q95qv6 caenorhabdi	623	37	50.7	243	2	Q4ZPX5_PSESY	Q4zpx5 pseudomonas
551	37	50.7	110	2	Q856U7_9CAUD	Q856u7 mycobacteri	624	37	50.7	243	2	Q75808_9HIV1	Q75808 human immun
552	37	50.7	110	2	Q8H246_VACMY	Q8h246 vaccinium m	625	37	50.7	243	2	Q75801_9HIV1	Q75801 human immun
553	37	50.7	111	2	Q86Z67_HETAN	Q86z67 heterobasid	626	37	50.7	243	2	Q58S24_9HIV1	Q58s24 human immun
554	37	50.7	111	2	Q86ZQ4_HETAN	Q86zq4 heterobasid	627	37	50.7	244	2	Q4V8S1_BRARE	Q4v8s1 brachydanio
555	37	50.7	111	2	Q86ZQ5_HETAN	Q86zq5 heterobasid	628	37	50.7	245	1	KDSB_FUSNN	Q8rfa8 fusobacteri
556	37	50.7	111	2	Q86ZQ6_HETAN	Q86zq6 heterobasid	629	37	50.7	245	2	Q6H0J6_9HOMO	Q6h0j6 leccinum cr
557	37	50.7	120	2	Q506S1_9ROSA	Q506s1 ficus bulle	630	37	50.7	245	2	Q6V776_HORSP	Q6v776 hordeum spo
558	37	50.7	120	2	Q506T1_9ROSA	Q506t1 ficus popen	631	37	50.7	245	2	Q6V785_HORSP	Q6v785 hordeum spo
559	37	50.7	120	2	Q506U4_9ROSA	Q506u4 ficus cf. t	632	37	50.7	245	2	Q7P7V1_FUSNV	Q7p7v1 fusobacteri
560	37	50.7	120	2	Q506U5_9ROSA	Q506u5 ficus insip	633	37	50.7	245	2	Q6QTJ9_9HIV1	Q6qtj9 human immun
561	37	50.7	120	2	Q506U6_9ROSA	Q506u6 ficus glabr	634	37	50.7	245	2	Q58S25_9HIV1	Q58s25 human immun
562	37	50.7	120	2	Q506U7_9ROSA	Q506u7 ficus yopon	635	37	50.7	247	2	Q75785_9HIV1	Q75785 human immun
563	37	50.7	120	2	Q506U8_9ROSA	Q506u8 ficus maxim	636	37	50.7	247	2	Q75807_9HIV1	Q75807 human immun
564	37	50.7	120	2	Q506U9_9ROSA	Q506u9 ficus dugan	637	37	50.7	247	2	Q75802_9HIV1	Q75802 human immun
565	37	50.7	120	2	Q506V0_9ROSA	Q506v0 ficus trigo	638	37	50.7	248	2	Q6EFI6_9HIV1	Q6efi6 human immun
566	37	50.7	120	2	Q506V1_9ROSA	Q506v1 ficus citri	639	37	50.7	249	2	Q6H0J5_9HOMO	Q6h0j5 leccinum ca
567	37	50.7	120	2	Q506V2_9ROSA	Q506v2 ficus parae	640	37	50.7	250	2	Q903C2_9HIV1	Q903c2 human immun
568	37	50.7	120	2	Q506V3_9ROSA	Q506v3 ficus colub	641	37	50.7	253	2	Q5MXH9_9HIV1	Q5mhx9 human immun
569	37	50.7	120	2	Q506V4_9ROSA	Q506v4 ficus costa	642	37	50.7	255	2	Q7ZKF3_9HIV1	Q7zkf3 human immun
570	37	50.7	120	2	Q506V5_9ROSA	Q506v5 ficus perfo	643	37	50.7	256	2	Q90LK9_9HIV1	Q90lk9 human immun
571	37	50.7	120	2	Q506V6_9ROSA	Q506v6 ficus pertu	644	37	50.7	259	2	Q56922_9HIV1	Q56922 human immun
572	37	50.7	120	2	Q506V7_9ROSA	Q506v7 ficus obtus	645	37	50.7	261	2	Q5RRV0_9HIV1	Q5rrv0 human immun
573	37	50.7	127	2	Q97401_9HIV1	Q97401 human immun	646	37	50.7	263	2	Q5LWV4_SILPO	Q5lvw4 silicibacte
574	37	50.7	127	2	Q97396_9HIV1	Q97396 human immun	647	37	50.7	266	2	Q6EFL4_9HIV1	Q6efl4 human immun
575	37	50.7	128	2	Q6G0X8_BARQU	Q6g0x8 bartonella	648	37	50.7	268	2	Q98LG1_RHILO	Q98lg1 rhizobium l
576	37	50.7	133	2	Q944T4_FRAAN	Q944t4 fragaria an	649	37	50.7	270	2	Q4SA81_TETNG	Q4sa81 tetraodon n
577	37	50.7	133	2	Q944T8_FRAAN	Q944t8 fragaria an	650	37	50.7	292	2	Q5GYG8_XANOR	Q5gyg8 xanthomonas
578	37	50.7	135	2	Q91Y39_RAT	Q91y39 rattus norv	651	37	50.7	292	2	Q72E10_DESVH	Q72el0 desulfovibr
579	37	50.7	136	2	Q5ZFK0_POPTN	Q5zfk0 populus tre	652	37	50.7	295	2	Q53I52_LUPAL	Q53i52 lupinus alb
580	37	50.7	138	2	Q82BS3_STRAW	Q82bs3 streptomyce	653	37	50.7	295	2	Q9EFG8_9HIV1	Q9efg8 human immun
581	37	50.7	139	2	Q9GSF9_HYDAT	Q9gsf9 hydra atten	654	37	50.7	296	2	Q4IBQ0_GIBZE	Q4ibq0 gibberella
582	37	50.7	139	2	Q9L2H5_STRCO	Q9l2h5 streptomyce	655	37	50.7	297	2	Q03307_CHELO	Q03307 chelodina l
583	37	50.7	140	2	Q5I5F3_9HOMO	Q5i5f3 leccinum sn	656	37	50.7	297	2	Q03335_9SAUR	Q03335 pelusios wi
584	37	50.7	140	2	Q82BN7_STRAW	Q82bn7 streptomyce	657	37	50.7	298	2	Q8J800_9HIV1	Q8j800 human immun
585	37	50.7	142	2	Q6YZA0_ORYSA	Q6yza0 oryza sativ	658	37	50.7	299	2	Q5RMX3_9HIV1	Q5rmx3 human immun
586	37	50.7	147	2	Q4UWE1_XANCP	Q4uwe1 xanthomonas	659	37	50.7	300	2	Q6R664_9HIV1	Q6r664 human immun
587	37	50.7	147	2	Q82Y81_NITEU	Q82y81 nitrosomona	660	37	50.7	300	2	Q5RRL4_9HIV1	Q5rrl4 human immun
588	37	50.7	147	2	Q8P7Q9_XANCP	Q8p7q9 xanthomonas	661	37	50.7	300	2	Q5RN14_9HIV1	Q5rn14 human immun
589	37	50.7	151	2	Q72880_9HIV1	Q72880 human immun	662	37	50.7	301	2	Q4NHT3_9MICC	Q4nht3 arthrobacte
590	37	50.7	152	2	Q8Q9F0_9HIV1	Q8q9f0 human immun	663	37	50.7	302	2	Q8KLC3_RHIET	Q8klc3 rhizobium e
591	37	50.7	152	2	Q8Q9F3_9HIV1	Q8q9f3 human immun	664	37	50.7	302	2	Q5RMI1_9HIV1	Q5rmil1 human immun
592	37	50.7	165	2	Q9A5T0_CAUCR	Q9a5t0 caulobacter	665	37	50.7	303	2	Q8J821_9HIV1	Q8j821 human immun
593	37	50.7	166	2	Q6TIB6_9HIV1	Q6tib6 human immun	666	37	50.7	304	2	Q5RRZ7_9HIV1	Q5rrz7 human immun
594	37	50.7	168	2	Q7XJU1_WHEAT	Q7xjj1 triticum ae	667	37	50.7	305	1	G3PC_HORVU	P08477 hordeum vul
595	37	50.7	168	2	Q9B596_CHILA	Q9b596 chinchilla	668	37	50.7	308	2	Q63NQ8_BURPS	Q63nq8 burkholderi
596	37	50.7	168	2	Q4TFP2_TETNG	Q4tfp2 tetraodon n	669	37	50.7	308	2	Q62AA0_BURMA	Q62aa0 burkholderi
597	37	50.7	173	2	Q8PLV3_XANAC	Q8plv3 xanthomonas	670	37	50.7	308	2	Q6R6F9_9HIV1	Q6r6f9 human immun
598	37	50.7	175	2	Q69WK4_ORYSA	Q69mk4 oryza sativ	671	37	50.7	313	2	Q8VNZ1_TREHY	Q8vnz1 treponema h
599	37	50.7	179	2	Q886G5_PSESM	Q886g5 pseudomonas	672	37	50.7	315	2	Q98ZH3_9HIV1	Q98zh3 human immun
600	37	50.7	182	2	Q9B598_CHILA	Q9b598 chinchilla	673	37	50.7	315	2	Q6WEM0_9STRA	Q6wew0 apodachlya
601	37	50.7	190	2	Q9SHP3_ARATH	Q9shp3 arabidopsis	674	37	50.7	318	2	Q76EG1_9EUGL	Q76eg1 eutreptiell
602	37	50.7	192	2	Q7NT62_CHRVO	Q7nt62 chromobacte	675	37	50.7	319	2	Q50A19_9HIV1	Q50ai9 human immun
603	37	50.7	192	2	Q955I9_9NEOB	Q955i9 indirana sp	676	37	50.7	319	2	Q98YH8_9HIV1	Q98yh8 human immun
604	37	50.7	193	2	Q5VJE4_9BURK	Q5vje4 brachymonas	677	37	50.7	320	2	Q98JB46_9HIV1	Q98jb46 human immun
605	37	50.7	195	2	Q6QJAL_COCIM	Q6qjal coccidioid	678	37	50.7	321	2	Q8JB46_9HIV1	Q8jb46 human immun
606	37	50.7	195	2	Q83J79_SHIFL	Q83j79 shigella fl	679	37	50.7	321	2	Q68Q60_9HIV1	Q68q60 human immun
607	37	50.7	197	2	Q9RP96_STRFR	Q9rp96 streptomyce	680	37	50.7	323	2	Q8XQW6_RALSO	Q8xqw6 ralstonia s
608	37	50.7	203	2	Q76EK6_9LAMI	Q76ek6 torenia hyb	681	37	50.7	324	2	Q6WPY7_9HIV1	Q6wpy7 human immun
609	37	50.7	206	2	Q622A8_CAEBR	Q622a8 caenorhabdi	682	37	50.7	328	2	Q9EDI0_9HIV1	Q9edi0 human immun
610	37	50.7	207	2	Q7XHP1_ORYSA	Q7xhpl1 oryza sativ	683	37	50.7	328	2	Q9ECQ1_9HIV1	Q9ecq1 human immun
611	37	50.7	217	2	Q6FEV9_ACIAAD	Q6fev9 acinetobact	684	37	50.7	328	2	Q9EEI3_9HIV1	Q9eei3 human immun
612	37	50.7	222	2	Q5RMI0_9HIV1	Q5rmi0 human immun	685	37	50.7	328	2	Q9EINI_9HIV1	Q9ein1 human immun
613	37	50.7	226	2	Q8Q3M7_9HIV1	Q8q3m7 human immun	686	37	50.7	328	2	Q9EFH1_9HIV1	Q9efh1 human immun
614	37	50.7	226	2	Q8Q3N4_9HIV1	Q8q3n4 human immun	687	37	50.7	328	2	Q9EE76_9HIV1	Q9ee76 human immun
615	37	50.7	229	2	Q9WKB0_9HIV1	Q9wkb0 human immun	688	37	50.7	328	2	Q9EEW3_9HIV1	Q9eew3 human immun

689	37	50.7	328	2	Q9ECX4_9HIV1	Q9ecx4 human immun	762	37	50.7	384	2	Q9J3Z1_9HIV1	Q9j3z1 human immun
690	37	50.7	328	2	Q9EEC5_9HIV1	Q9eec5 human immun	763	37	50.7	386	1	ETR1_CANTR	Q8wzm3 candida tro
691	37	50.7	328	2	Q9EKL1_9HIV1	Q9ekl1 human immun	764	37	50.7	386	1	ETR2_CANTR	Q8wzm4 candida tro
692	37	50.7	328	2	Q9ECJ5_9HIV1	Q9ecj5 human immun	765	37	50.7	387	2	Q90GP0_9HIV1	Q90gp0 human immun
693	37	50.7	328	2	Q9EHX6_9HIV1	Q9ehx6 human immun	766	37	50.7	391	2	Q9S2B3_STRCO	Q9s2b3 streptomyce
694	37	50.7	328	2	Q9EL92_9HIV1	Q9el92 human immun	767	37	50.7	391	2	Q98YZ8_9HIV1	Q98yz8 human immun
695	37	50.7	328	2	Q9EDJ2_9HIV1	Q9edj2 human immun	768	37	50.7	392	2	O36575_9HIV1	O36575 human immun
696	37	50.7	328	2	Q9EHC6_9HIV1	Q9ehc6 human immun	769	37	50.7	393	2	O8PT05_METWA	O8pt05 methanosarc
697	37	50.7	328	2	Q9ECF3_9HIV1	Q9ecf3 human immun	770	37	50.7	395	2	Q5FQ25_GLUOX	Q5fq25 gluconobact
698	37	50.7	328	2	Q9EGW3_9HIV1	Q9egw3 human immun	771	37	50.7	398	2	Q4FDW6_DROXI	Q4fdw6 drosophila
699	37	50.7	329	2	Q904D4_9HIV1	Q904d4 human immun	772	37	50.7	398	2	Q5M7R8_XENTR	Q5m7r8 xenopus tro
700	37	50.7	329	2	Q8QBR7_9HIV1	Q8qbr7 human immun	773	37	50.7	401	2	Q7VVG2_BORPE	Q7vvg2 bordetella
701	37	50.7	330	2	Q50AS4_9HIV1	Q50as4 human immun	774	37	50.7	401	2	Q7WB96_BORPA	Q7wb96 bordetella
702	37	50.7	331	1	G3P_ASHGO	Q757i2 ashbya goss	775	37	50.7	401	2	Q7WMR4_BORBR	Q7wmr4 bordetella
703	37	50.7	331	2	Q5PY03_MUSAC	Q5py03 musa acumin	776	37	50.7	401	2	Q90IW3_9HIV1	Q90iw3 human immun
704	37	50.7	331	2	Q4JU09_CORJK	Q4ju09 corynebacte	777	37	50.7	405	2	Q5LQ49_SILPO	Q5lq49 silicibacte
705	37	50.7	332	2	Q4WDX7_ASPFU	Q4wdx7 aspergillus	778	37	50.7	406	2	Q4HCE3_9DEIO	Q4hce3 deinococcus
706	37	50.7	332	2	Q8L6S7_9EUGL	Q8l6s7 euglena sp.	779	37	50.7	407	2	Q9SA29_ARATH	Q9sa29 arabidopsis
707	37	50.7	332	2	Q8Q9H6_9HIV1	Q8q9h6 human immun	780	37	50.7	409	2	Q90MT8_9HIV1	Q90mt8 human immun
708	37	50.7	332	2	Q8Q9J8_9HIV1	Q8q9j8 human immun	781	37	50.7	410	2	Q4LL06_9BURK	Q4ll06 burkholderi
709	37	50.7	332	2	Q8Q9H4_9HIV1	Q8q9h4 human immun	782	37	50.7	411	2	Q6H703_ORYSA	Q6h703 oryza sativ
710	37	50.7	332	2	Q8Q9H3_9HIV1	Q8q9h3 human immun	783	37	50.7	413	2	Q6LGY3_PHOPR	Q6lgy3 photobacter
711	37	50.7	332	2	Q8Q9I9_9HIV1	Q8q9i9 human immun	784	37	50.7	414	2	Q87FL6_VIBPA	Q87fl6 vibrio para
712	37	50.7	332	2	Q8Q9I2_9HIV1	Q8q9i2 human immun	785	37	50.7	415	2	Q655W2_ORYSA	Q655w2 oryza sativ
713	37	50.7	332	2	Q7ZM79_9HIV1	Q7zm79 human immun	786	37	50.7	416	2	Q8DHK5_SYNEL	Q8dhk5 synechococc
714	37	50.7	332	2	Q7ZM65_9HIV1	Q7zm65 human immun	787	37	50.7	417	2	Q8AEC2_9HIV1	Q8aec2 human immun
715	37	50.7	332	2	Q7ZM00_9HIV1	Q7zm00 human immun	788	37	50.7	419	2	Q8VWN9_CAPAN	Q8vwn9 capsicum an
716	37	50.7	332	2	Q7ZLZ8_9HIV1	Q7z1z8 human immun	789	37	50.7	420	2	Q5E924_ARATH	Q5e924 arabidopsis
717	37	50.7	332	2	Q7ZLY1_9HIV1	Q7z1y1 human immun	790	37	50.7	420	2	Q84WR0_ARATH	Q84wr0 arabidopsis
718	37	50.7	333	2	Q9FVP0_ASCNO	Q9fvp0 ascophyllum	791	37	50.7	420	2	Q8L8W7_ARATH	Q8l8w7 arabidopsis
719	37	50.7	333	2	Q8JA16_9HIV1	Q8ja16 human immun	792	37	50.7	422	2	Q9SAJ6_ARATH	Q9saj6 arabidopsis
720	37	50.7	335	1	G3P1_TRIKO	P17729 trichoderma	793	37	50.7	423	2	Q9QCY1_9HIV1	Q9qcy1 human immun
721	37	50.7	335	2	Q8QAQ1_9HIV1	Q8qaq1 human immun	794	37	50.7	423	2	Q8AHP1_9HIV1	Q8ahp1 human immun
722	37	50.7	335	2	Q8QAP6_9HIV1	Q8qap6 human immun	795	37	50.7	423	2	Q8AH66_9HIV1	Q8ah66 human immun
723	37	50.7	335	2	Q8QAQ8_9HIV1	Q8qaq8 human immun	796	37	50.7	425	2	Q8AH61_9HIV1	Q8ah61 human immun
724	37	50.7	335	2	Q533H4_9HIV1	Q533h4 human immun	797	37	50.7	426	2	Q854V2_9CAUD	Q854v2 mycobacteri
725	37	50.7	337	1	G3PC_CRAPL	Q42671 craterostig	798	37	50.7	426	2	Q4VZ57_9HIV1	Q4vz57 human immun
726	37	50.7	337	1	G3PC_ORYSA	Q42977 oryza sativ	799	37	50.7	433	2	Q37264_PINSY	Q37264 pinus sylve
727	37	50.7	337	1	G3PE_MAIZE	Q43247 zea mays (m	800	37	50.7	433	2	Q37265_PINSY	Q37265 pinus sylve
728	37	50.7	337	1	G3PX_HORVU	P26517 hordeum vul	801	37	50.7	434	2	Q4JK01_9HIV1	Q4jk01 human immun
729	37	50.7	337	2	Q43359_MAIZE	Q43359 zea mays (m	802	37	50.7	434	2	Q5D836_9HIV1	Q5d836 human immun
730	37	50.7	337	2	Q6ZK60_ORYSA	Q6zk60 oryza sativ	803	37	50.7	438	2	Q9AA92_CAUCR	Q9aa92 caulobacter
731	37	50.7	337	2	Q7FAH2_ORYSA	Q7fah2 oryza sativ	804	37	50.7	443	2	Q4UR06_XANCP	Q4ur06 xanthomonas
732	37	50.7	338	1	G3PC_PEA	P34922 pisum sativ	805	37	50.7	443	2	Q8P330_XANCP	Q8p330 xanthomonas
733	37	50.7	338	1	G3PC_RANAC	P26521 ranunculus	806	37	50.7	452	2	Q9YH86_CHICK	Q9yh86 gallus gall
734	37	50.7	340	2	Q50BB8_9HIV1	Q50bb8 human immun	807	37	50.7	460	2	Q93P09_PROMP	Q93p09 prochloroco
735	37	50.7	341	1	G3PC_CHLRE	P49644 chlamydomon	808	37	50.7	461	2	Q60CR2_METCA	Q60cr2 methylococc
736	37	50.7	341	1	G3PC_MAGLI	P26518 magnolia li	809	37	50.7	465	2	Q7V6H9_PROMM	Q7v6h9 prochloroco
737	37	50.7	341	2	Q8ZFT1_YERPE	Q8zft1 yersinia pe	810	37	50.7	477	2	Q88GH3_PSEPK	Q88gh3 pseudomonas
738	37	50.7	341	2	Q669L7_YERPS	Q669l7 yersinia ps	811	37	50.7	480	2	O22813_ARATH	O22813 arabidopsis
739	37	50.7	341	2	Q7NNT5_GLOVI	Q7nnt5 gloeobacter	812	37	50.7	481	2	Q8L117_RHOSH	Q8l117 rhodobacter
740	37	50.7	344	2	Q7VVKH5_HAEDU	Q7vkh5 haemophilus	813	37	50.7	486	2	Q4WM98_ASPFU	Q4wm98 aspergillus
741	37	50.7	345	2	Q6DA67_ERWCT	Q6da67 erwinia car	814	37	50.7	486	2	Q9SGQ0_ARATH	Q9sgq0 arabidopsis
742	37	50.7	349	2	Q5R0S2_IDILO	Q5r0s2 idiomarina	815	37	50.7	486	2	Q5H1W5_XANOR	Q5hiw5 xanthomonas
743	37	50.7	349	2	Q90M47_9HIV1	Q90m47 human immun	816	37	50.7	491	2	Q7MCZ2_VIBVY	Q7mcz2 vibrio vuln
744	37	50.7	351	2	Q8Q9K2_9HIV1	Q8q9k2 human immun	817	37	50.7	491	2	Q5L085_GEOKA	Q5l085 geobacillus
745	37	50.7	351	2	Q8Q9K2_9HIV1	Q8q9k2 human immun	818	37	50.7	499	2	Q6RT41_9HIV1	Q6rt41 human immun
746	37	50.7	352	2	Q9I763_PSEAE	Q9i763 pseudomonas	819	37	50.7	499	2	Q67D18_9HIV1	Q67d18 human immun
747	37	50.7	353	2	Q43311_EUGGR	Q43311 euglena gra	820	37	50.7	499	2	Q67D97_9HIV1	Q67d97 human immun
748	37	50.7	356	2	Q8AEA6_9HIV1	Q8aea6 human immun	821	37	50.7	499	2	Q67D81_9HIV1	Q67d81 human immun
749	37	50.7	360	2	Q90KA3_9HIV1	Q90ka3 human immun	822	37	50.7	499	2	Q67CP5_9HIV1	Q67cp5 human immun
750	37	50.7	361	2	Q5GGD9_9HIV1	Q5ggd9 human immun	823	37	50.7	499	2	Q67CB1_9HIV1	Q67cb1 human immun
751	37	50.7	362	1	MRAY_AZOSE	Q5p6z3 azoarcus sp	824	37	50.7	499	2	Q67CA2_9HIV1	Q67ca2 human immun
752	37	50.7	363	2	Q8D0Q7_YERPE	Q8d0q7 yersinia pe	825	37	50.7	499	2	Q5I1W4_9HIV1	Q5ilw4 human immun
753	37	50.7	364	1	MRAW_BORBR	Q7wfr5 bordetella	826	37	50.7	499	2	Q5I1T3_9HIV1	Q5ilt3 human immun
754	37	50.7	364	1	MRAW_BORPA	Q7w4a7 bordetella	827	37	50.7	505	2	Q7UGI8_RHOBA	Q7ugi8 rhodopirell
755	37	50.7	364	1	MRAW_BORPE	Q7vup6 bordetella	828	37	50.7	511	2	Q83C84_COXBU	Q83c84 coxiella bu
756	37	50.7	365	2	Q65876_PINSY	Q65876 pinus sylve	829	37	50.7	513	2	Q7T2M9_CARAU	Q7t2m9 carassius a
757	37	50.7	366	2	Q6LNZ5_PHOPR	Q6lnz5 photobacter	830	37	50.7	524	2	Q9IDL5_9HIV1	Q9idl5 human immun
758	37	50.7	375	2	Q6QW16_AZOBR	Q6qw16 azospirillu	831	37	50.7	531	2	Q4Q9J5_LEIMA	Q4q9j5 leishmania
759	37	50.7	380	2	Q7U4M1_SYNPX	Q7u4m1 synechococc	832	37	50.7	531	2	Q63V30_BURPS	Q63v30 burkholderi
760	37	50.7	382	2	Q8U337_PYRFU	Q8u337 pyrococcus	833	37	50.7	531	2	Q62JL9_BURMA	Q62jl9 burkholderi
761	37	50.7	383	2	Q53NA8_ORYSA	Q53na8 oryza sativ	834	37	50.7	547	2	Q9X8S7_STRCO	Q9x8s7 streptomyce

835	37	50.7	552	2	Q7NX80_CHRVO	Q7nx80_chromobacte	908	37	50.7	1043	2	Q90DD2_SIVCZ	Q90dd2_chimpanzee
836	37	50.7	556	2	Q6ONX9_CAEBR	Q60nx9_caenorhabdi	909	37	50.7	1055	2	Q89928_9HIV2	Q89928_human immun
837	37	50.7	559	2	Q8UTW0_9HIV1	Q8utw0_human immun	910	37	50.7	1061	2	Q76634_9HIV2	Q76634_human immun
838	37	50.7	560	2	Q5DNK6_9HIV1	Q5dnk6_human immun	911	37	50.7	1064	2	Q8UTV7_9HIV2	Q8utv7_human immun
839	37	50.7	562	2	Q99B10_9HIV1	Q99b10_human immun	912	37	50.7	1086	2	Q7UIK4_RHOBA	Q7uik4_rhodopirell
840	37	50.7	564	2	Q9J4V0_9HIV1	Q9j4v0_human immun	913	37	50.7	1091	1	RHG07_RAT	Q63744_r rho-gtpas
841	37	50.7	569	2	Q9MA84_ARATH	Q9ma84_arabidopsis	914	37	50.7	1116	2	Q4V0E3_XANCP	Q4v0e3_xanthomonas
842	37	50.7	574	1	SYE_AERPE	Q9y9hi_aeropyrum p	915	37	50.7	1116	2	Q8PE49_XANCP	Q8pe49_xanthomonas
843	37	50.7	577	2	Q9IN28_9HIV1	Q9in28_human immun	916	37	50.7	1116	2	Q8PR14_XANAC	Q8pr14_xanthomonas
844	37	50.7	586	2	Q70ER5_9HIV1	Q70er5_human immun	917	37	50.7	1123	2	Q66933_9RETR	Q66933_feline immu
845	37	50.7	596	2	Q9M7R8_ACHBI	Q9m7r8_achlya bise	918	37	50.7	1168	2	Q7SH52_NEUCR	Q7sh52_neurospora
846	37	50.7	597	2	Q98KC8_RHILO	Q98kc8_rhizobium l	919	37	50.7	1170	2	Q4INJ1_GIBZE	Q4inj1_gibberella
847	37	50.7	614	2	Q9LE93_PHATR	Q9le93_phaeodactyl	920	37	50.7	1233	1	VL1_REOVD	Q4injl_reovirus ty
848	37	50.7	615	2	Q9M7R5_ODOSI	Q9m7r5_odontella s	921	37	50.7	1275	2	Q9WAB0_REOVD	Q9wab0_reovirus ty
849	37	50.7	619	2	Q4FZ41_LEIMA	Q4fz41_leishmania	922	37	50.7	1275	2	Q9WAB2_REOVL	Q9wab2_reovirus ty
850	37	50.7	620	2	Q5TSW6_ANOGA	Q5tsw6_anopheles g	923	37	50.7	1275	2	Q9WAB1_REOVJ	Q9wab1_reovirus ty
851	37	50.7	630	2	Q5AVJ2_EMENI	Q5avj2_aspergillus	924	37	50.7	1357	2	Q5AXV0_EMENI	Q5axv0_aspergillus
852	37	50.7	657	2	Q6JSB8_9HIV1	Q6jsb8_human immun	925	37	50.7	1391	2	Q4N6K3_THEPA	Q4n6k3_theileria p
853	37	50.7	657	2	Q6JSB7_9HIV1	Q6jsb7_human immun	926	37	50.7	1413	2	Q4UJ10_THEAN	Q4uj10_theileria a
854	37	50.7	657	2	Q6JSB6_9HIV1	Q6jsb6_human immun	927	37	50.7	1435	2	Q41791_9HIV1	Q41791_human immun
855	37	50.7	657	2	Q6JSB5_9HIV1	Q6jsb5_human immun	928	37	50.7	1663	2	Q88L55_PSEPK	Q88l55_pseudomonas
856	37	50.7	657	2	Q6JSB4_9HIV1	Q6jsb4_human immun	929	37	50.7	1955	2	Q4PE01_USTMA	Q4pe01_ustilago ma
857	37	50.7	657	2	Q6JSA7_9HIV1	Q6jsa7_human immun	930	37	50.7	2108	2	Q7VEZ3_MYCBO	Q7vez3_mycobacteri
858	37	50.7	657	2	Q6JSA6_9HIV1	Q6jsa6_human immun	931	37	50.7	2108	2	O53901_MYCTU	O53901_mycobacteri
859	37	50.7	664	2	Q9K9Z0_BACHD	Q9k9z0_bacillus ha	932	37	50.7	2229	2	Q9EWP5_STRCO	Q9ewp5_streptomyce
860	37	50.7	669	2	Q526R7_MAGGR	Q526r7_magnaporth	933	37	50.7	2307	2	Q612L3_CAEBR	Q612l3_caenorhabdi
861	37	50.7	679	2	Q51U94_MAGGR	Q51u94_magnaporth	934	37	50.7	2329	1	PRP3_CAEEL	P34369_caenorhabdi
862	37	50.7	700	2	Q7Q9W9_ANOGA	Q7q9w9_anopheles g	935	37	50.7	2363	2	O14187_SCHPO	O14187_schizosacch
863	37	50.7	706	2	Q7Z2M3_HUMAN	Q7z2m3_homo sapien	936	37	50.7	3165	2	O5AR63_EMENI	O5ar63_aspergillus
864	37	50.7	710	2	Q8RY11_ARATH	Q8ry11_arabidopsis	937	36.5	50.0	88	2	Q6YS15_ORYSA	Q6ys15_oryza sativ
865	37	50.7	712	2	Q7MUI4_PORGI	Q7mui4_porphyromon	938	36.5	50.0	133	2	Q92RI6_RHIME	Q92ri6_rhizobium m
866	37	50.7	715	2	Q9U6D9_DROME	Q9u6d9_drosophila	939	36.5	50.0	305	1	UVSE_DEIRA	Q9rte6_deinococcus
867	37	50.7	715	2	Q9UB09_DROME	Q9ub09_drosophila	940	36.5	50.0	339	2	Q4J2R2_AZOVI	Q4j2r2_azotobacter
868	37	50.7	715	2	Q96689_DROME	Q96689_drosophila	941	36.5	50.0	423	2	Q4LQN8_9BURK	Q4lqn8_burkholderi
869	37	50.7	724	2	Q5CPL7_CRYPV	Q5cpl7_cryptospori	942	36.5	50.0	502	2	Q9A4M5_CAUCR	Q9a4m5_caulobacter
870	37	50.7	724	2	Q5CK52_CRYHO	Q5ck52_cryptospori	943	36.5	50.0	529	2	Q8PQZ7_XANAC	Q8pqz7_xanthomonas
871	37	50.7	730	2	Q4PBC2_USTWA	Q4pbc2_ustilago ma	944	36.5	50.0	2221	2	Q5AV42_EMENI	Q5av42_aspergillus
872	37	50.7	749	2	Q5FR28_GLUOX	Q5fr28_gluconobact	945	36	49.3	60	1	RL28_BACST	P23374_bacillus st
873	37	50.7	755	2	Q5SR90_CRYNE	Q5sr90_cryptococcu	946	36	49.3	61	1	RL28_GEOKA	Q510r4_geobacillus
874	37	50.7	755	2	Q5KEY0_CRYNE	Q5key0_cryptococcu	947	36	49.3	61	1	RL28_LACJO	Q74in6_lactobacill
875	37	50.7	755	2	Q9I2P9_PSEAE	Q9i2p9_pseudomonas	948	36	49.3	77	2	Q8HZ44_PAPHA	Q8hz44_papio hamad
876	37	50.7	762	2	Q88126_PSESM	Q881z6_pseudomonas	949	36	49.3	81	2	Q9NDL5_9CNID	Q9ndl5_tima formos
877	37	50.7	777	1	METE_CAUCR	Q9aaw1_caulobacter	950	36	49.3	81	2	Q9NDL8_HYDEC	Q9ndl8_hydractinia
878	37	50.7	785	2	Q7NPF6_GLOVI	Q7npf6_gloeobacter	951	36	49.3	81	2	Q9NDL9_9CNID	Q9ndl9_eirene sp.
879	37	50.7	804	2	Q4PG05_USTMA	Q4pg05_ustilago ma	952	36	49.3	85	2	Q64GR5_9HIV1	Q64gr5_human immun
880	37	50.7	825	1	SE5_RAT	Q63003_rattus norv	953	36	49.3	93	1	Y2256_XANCP	Q4uvk1_xanthomonas
881	37	50.7	829	2	Q4RX45_TETNG	Q4rx45_tetraodon n	954	36	49.3	93	2	Q4UVK1_XANCP	Q4uvk1_xanthomonas
882	37	50.7	838	2	Q7MRT1_WOLSU	Q7mrt1_wolinella s	955	36	49.3	95	1	Y2359_XANAC	Q8pk18_xanthomonas
883	37	50.7	853	2	Q8Q5Q4_9RETR	Q8q5q4_feline immu	956	36	49.3	97	2	Q8FPC4_COREF	Q8fpc4_corynebacte
884	37	50.7	854	2	Q6J4Y9_9RETR	Q6j4y9_feline immu	957	36	49.3	98	2	Q599U7_9BACT	Q599u7_uncultured
885	37	50.7	858	2	Q782U1_9ALPH	Q782u1_gallid herp	958	36	49.3	102	2	Q33137_MYCLE	Q33137_mycobacteri
886	37	50.7	858	2	Q9QTD8_9ALPH	Q9qtd8_marek's dis	959	36	49.3	104	2	Q92U58_RHIME	Q92u58_rhizobium m
887	37	50.7	890	2	Q4URT7_XANCP	Q4urt7_xanthomonas	960	36	49.3	117	2	Q82C58_STRAW	Q82c58_streptomyce
888	37	50.7	890	2	Q7U6X4_SYNPX	Q7u6x4_synechococc	961	36	49.3	119	2	Q8N4D3_HUMAN	Q8n4d3_homo sapien
889	37	50.7	890	2	Q8PBR5_XANCP	Q8pb5_xanthomonas	962	36	49.3	120	2	Q6ZSJ1_HUMAN	Q6zsj1_homo sapien
890	37	50.7	890	2	Q5H6H0_XANOR	Q5h6h0_xanthomonas	963	36	49.3	121	2	Q5V089_HALMA	Q5v089_haloarcula
891	37	50.7	912	2	Q9Q6Z2_9HIV1	Q9q6z2_human immun	964	36	49.3	126	2	Q7QQ70_GIALA	Q7qq70_giardia lam
892	37	50.7	931	2	Q7WY21_PSEAE	Q7wy21_pseudomonas	965	36	49.3	127	2	Q6G5H7_BARHE	Q6g5h7_bartonella
893	37	50.7	940	2	Q7UQT9_RHOBA	Q7uqt9_rhodopirell	966	36	49.3	128	1	NBSM_HUMAN	Q95168_homo sapien
894	37	50.7	956	1	DPO1_DEIRA	P52027_deinococcus	967	36	49.3	129	2	Q5R6P3_PONPY	Q5r6p3_pongo pygma
895	37	50.7	999	2	Q8UTG9_9HIV1	Q8utg9_human immun	968	36	49.3	131	2	Q9CVZ3_MOUSE	Q9cvz3_mus musculu
896	37	50.7	1000	2	Q6S7Y7_9HIV1	Q6s7y7_human immun	969	36	49.3	132	2	Q6MPB9_BDEBA	Q6mpb9_bdellovibri
897	37	50.7	1002	2	Q9QST9_9HIV1	Q9qst9_human immun	970	36	49.3	133	1	EPPI_HUMAN	Q95925_homo sapien
898	37	50.7	1002	2	Q4QX95_9HIV1	Q4qx95_human immun	971	36	49.3	133	1	EPPI_MACMU	Q9bdl1_macaca mula
899	37	50.7	1003	2	Q6PR23_9HIV1	Q6pr23_human immun	972	36	49.3	133	2	Q8HZ45_PAPPA	Q8hz45_papio papio
900	37	50.7	1003	2	Q5I9G4_9HIV1	Q5i9g4_human immun	973	36	49.3	134	2	Q5LF77_BACFN	Q5lf77_bacteroides
901	37	50.7	1006	2	Q5U8G6_9HIV1	Q5u8g6_human immun	974	36	49.3	135	2	Q8TN98_METAC	Q8tn98_methanosarc
902	37	50.7	1006	2	Q596M2_9HIV1	Q596m2_human immun	975	36	49.3	135	2	Q6FSL6_CANGA	Q6fsl6_candida gla
903	37	50.7	1007	2	Q66QF6_9HIV1	Q66qf6_human immun	976	36	49.3	142	2	Q8CY09_ECOL6	Q8cy09_escherichia
904	37	50.7	1009	1	POL_SIVGB	P22382_simian immu	977	36	49.3	152	2	Q82X11_NITEU	Q82xi1_nitrosomona
905	37	50.7	1009	2	Q9NKU7_LEIMA	Q9nku7_leishmania	978	36	49.3	153	2	Q9CAH9_ARATH	Q9cah9_arabidopsis
906	37	50.7	1016	2	Q8JAI0_SIVCZ	Q8jai0_chimpanzee	979	36	49.3	155	2	Q4IF93_GIBZE	Q4if93_gibberella
907	37	50.7	1039	2	Q87108_SIVCZ	Q87108_chimpanzee	980	36	49.3	156	2	Q76KB7_NYMAL	Q76kb7_nymphaea al

981 36 49.3 163 2 Q749N7_GEOSL Q749n7 geobacter s
982 36 49.3 172 2 Q67FF4_9HIV1 Q67ff4 human immun
983 36 49.3 176 2 Q656A5_ORYSA Q656a5 oryza sativ
984 36 49.3 176 2 Q6SGI2_9BACT Q6sg12 uncultured
985 36 49.3 176 2 Q6SHI2_9BACT Q6shi2 uncultured
986 36 49.3 179 2 Q8XSL7_RALSO Q8xsl7 ralstonia s
987 36 49.3 183 2 Q62KY8_BURMA Q62ky8 burkholderi
988 36 49.3 185 2 Q8CME3_SHEON Q8cme3 shewanella
989 36 49.3 186 2 Q703F2_LINUS Q703f2 linum usita
990 36 49.3 188 2 Q9PHI6_XYLFA Q9phi6 xylella fas
991 36 49.3 189 2 Q8CZ94_STRR6 Q8cz94 streptococc
992 36 49.3 189 2 Q97T07_STRPN Q97t07 streptococc
993 36 49.3 193 2 Q7NVI3_CHRVO Q7nvi3 chromobacte
994 36 49.3 193 2 Q4VTU6_9SAUR Q4vtu6 gymnodactyl
995 36 49.3 194 2 Q5SJF1_THET8 Q5sjf1 thermus the
996 36 49.3 197 2 Q92NS0_RHIME Q92ns0 rhizobium m
997 36 49.3 203 2 Q4VTU5_9SAUR Q4vtu5 gymnodactyl
998 36 49.3 204 1 THIE_RH1ET Q34294 rhizobium e
999 36 49.3 204 2 Q8Y1J0_RALSO Q8ylj0 ralstonia s
1000 36 49.3 207 2 Q61YT1_CAEBR Q61yt1 caenorhabdi

ALIGNMENTS

RESULT 1
Q5YWJ5_NOCFA PRELIMINARY; PRT; 53 AA.
ID Q5YWJ5_NOCFA PRELIMINARY; PRT; 53 AA.
AC Q5YWJ5;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=nfa25990;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD57446.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 53 AA; 5806 MW; 47B08B1F9BA4FCE6 CRC64;

Query Match 68.5%; Score 50; DB 2; Length 53;
Best Local Similarity 64.3%; Pred. No. 0.68;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
: || ||||| |
Db 31 VHGPQCRQWLAAA 44

RESULT 2
Q742B3_MYCPA PRELIMINARY; PRT; 302 AA.
ID Q742B3_MYCPA PRELIMINARY; PRT; 302 AA.
AC Q742B3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Galu.
GN Name=galu; OrderedLocusNames=NAP0924;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017230; AAS03241.1; -; Genomic DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;

Query Match 68.5%; Score 50; DB 2; Length 302;
Best Local Similarity 81.8%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
|| ||||| |
Db 286 GPDRLRWLVAR 296

RESULT 3
CBBR_XANFL STANDARD; PRT; 333 AA.
ID CBBR_XANFL STANDARD; PRT; 333 AA.
AC P25545;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE HTH-type transcriptional regulator cbbR (RuBisCO operon
transcriptional regulator).
GN Name=cbbR; Synonyms=cfxO;
OS Xanthobacter flavus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Xanthobacteraceae; Xanthobacter.
OX NCBI_TaxID=281;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H4-14;
RX MEDLINE=94012468; PubMed=8407781;
RA van den Bergh E., Dijkhuizen L., Meijer W.G.;
RT "CbbR, a LysR-type transcriptional activator, is required for
expression of the autotrophic CO2 fixation enzymes of Xanthobacter
flavus.";
RL J. Bacteriol. 175:6097-6104(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-150.
RC STRAIN=H4-14;
RX MEDLINE=91172133; PubMed=1900916; DOI=10.1007/BF00269865;
RA Meijer W.G., Arnberg A.C., Enequist H.G., Terpstra P., Lidstrom M.E.,
RA Dijkhuizen L.;
RT "Identification and organization of carbon dioxide fixation genes in
Xanthobacter flavus H4-14.";
RL Mol. Gen. Genet. 225:320-330(1991).
CC -!- FUNCTION: Transcriptional activator for the cbb operon (cbbLSXFP)
for RuBisCO and other Calvin cycle genes. Binds specifically to
two binding sites in the cbbR-cbbL intergenic region.
CC -!- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; Z22705; CAA80406.1; -; Genomic DNA.
DR EMBL; X17252; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; A36925; A36925.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.

DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR01590; HTHFIS.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
FT DOMAIN 5 62 HTH_LYSR-type.
FT DNA BIND 22 41 H-T-H motif (By similarity).
SQ SEQUENCE 333 AA; 36004 MW; 9B375B4FB2D1EE73 CRC64;

Query Match 67.8%; Score 49.5; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 IEG-PTLRQWLAAAR 14
:|||:|||||||
Db 264 VEGLPVVRQWLAVRA 278

RESULT 4
Q9RKMS_STRCO PRELIMINARY; PRT; 319 AA.
AC Q9RKMS5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MerR family transcriptional regulator.
GN OrderedLocusNames=SCO4102; ORFNames=SCD17.06c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cergeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939118; CAB56383.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS50937; HTH_MERR_2; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;

Query Match 67.1%; Score 49; DB 2; Length 319;
Best Local Similarity 66.7%; Pred. No. 6.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAAAR 13
:|||:|||||
Db 258 DGPRLREWLAGR 269

RESULT 5
Q9L8D4_POLCB PRELIMINARY; PRT; 607 AA.
ID Q9L8D4_POLCB PRELIMINARY;

AC Q9L8D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Polyangium cellulosum (Sorangium cellulosum).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Sorangineae; Polyangiaceae; Polyangium.
OX NCBI_TaxID=56;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=So ce90;
RX MEDLINE=20130945; PubMed=10662695; DOI=10.1016/S1074-5521(00)00075-2;
RA Molnar I., Schupp T., Ono M., Zirkle R.E., Milnamow M.,
RA Nowak-Thompson B., Engel N., Toupet C., Stratmann A., Cyr D.D.,
RA Gorlach J., Mayo J.M., Hu A., Goff S., Schmid J., Ligon J.M.;
RT "The biosynthetic gene cluster for the microtubule-stabilizing agents
epothilones A and B from Sorangium cellulosum So ce90.";
RL Chem. Biol. 7:97-109(2000).
DR EMBL; AF210843; AAF26904.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 607 AA; 66326 MW; F113CA299B25048E CRC64;

Query Match 65.8%; Score 48; DB 2; Length 607;
Best Local Similarity 61.5%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 13
:|||:|||||
Db 96 VDGPAVVRWLAAAR 108

RESULT 6
Q4KS46_9VIRU PRELIMINARY; PRT; 815 AA.
ID Q4KS46_9VIRU PRELIMINARY;
AC Q4KS46;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tyrosine kinase.
OS Orange-spotted grouper iridovirus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=322017;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhou S.Y., Lv L., Chen C., Weng S.P., Chan S.M., He J.G.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY894343; AAX82420.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 815 AA; 93153 MW; 120C2FB59703F8C4 CRC64;

Query Match 65.8%; Score 48; DB 2; Length 815;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAAR 14
:|||||:|:|
Db 500 VQGPTLAQWICSTA 513

RESULT 7
Q8QUJ6_9VIRU PRELIMINARY; PRT; 941 AA.
ID Q8QUJ6_9VIRU PRELIMINARY;
AC Q8QUJ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF114L.
OS Infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Megalocytivirus.
OX NCBI_TaxID=180170;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus.";
RL Virology 291:126-139(2001).
DR EMBL; AF371960; AAL98838.1; -; Genomic_DNA.
SQ SEQUENCE 941 AA; 106703 MW; EB663998C7F6CE83 CRC64;

Query Match 65.8%; Score 48; DB 2; Length 941;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:||||| :|
Db 581 VQGPTLAQWICSTA 594

RESULT 8
Q8ZGS7 YERPE PRELIMINARY; PRT; 296 AA.
AC Q8ZGS7; Q74WE0; Q7CH89;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative membrane protein (Putative transmembrane protein).
GN Name=thaT4; OrderedLocusNames=YP0934, YPO1203, Y2985;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";
RL DNA Res. 11:179-197(2004).
DR EMBL; AJ414147; CAC90042.1; -; Genomic DNA.
DR EMBL; AB013900; AAM86536.1; -; Genomic_DNA.
DR EMBL; AE017130; AAS61189.1; -; Genomic_DNA.
DR PIR; AG0147; AG0147.
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 296 AA; 31378 MW; 45947413DCD54CFF CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PTLRQWLAARA 14
||||| :|
Db 66 PTLRQWAAASA 76

RESULT 9
Q66D06 YERPS PRELIMINARY; PRT; 296 AA.
AC Q66D06;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative drug/metabolite (DME family) efflux pump precursor.
GN OrderedLocusNames=YPTB1243;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH20483.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 25 Potential.
SQ SEQUENCE 296 AA; 31407 MW; 4D3E486D32DBAC11 CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PTLRQWLAARA 14
||||| :|
Db 66 PTLRQWAAASA 76

RESULT 10
Q7D906 MYCTU PRELIMINARY; PRT; 306 AA.
AC Q7D906;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9).
GN Name=gau; OrderedLocusNames=MT1022;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK45269.1; -; Genomic_DNA.
DR TIGR; MT1022; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0003983; F:UTP-glucose-1-phosphate uridylyltransferase. . .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32406 MW; 880D3BB88CB0A3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
|||:|:|
Db 290 GPDLRRLVAR 300

RESULT 11
O05576_MYCTU
ID O05576_MYCTU PRELIMINARY; PRT; 306 AA.
AC O05576;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PROBABLE UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE)
DE (EC 2.7.7.9).
DE Name=gau; OrderedLocusNames=Rv0993;
GN Mycobacterium tuberculosis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaita F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842575; CAB08153.1; -; Genomic_DNA.
DR PIR; D70601; D70601.
DR Tuberculist; Rv0993; -.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32378 MW; 24C2387443B0A3E8 CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 GPTLRQWLAAR 13
|||:|:|
Db 290 GPDLRRLVAR 300

RESULT 12
Q7U0W3_MYCBO
ID Q7U0W3_MYCBO PRELIMINARY; PRT; 306 AA.
AC Q7U0W3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE)
DE (EC 2.7.7.9).
DE Name=gau; OrderedLocusNames=Mb1020;
GN Mycobacterium bovis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248337; CAD93881.1; -; Genomic_DNA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32406 MW; 880D3BB88CB0A3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
|||:|:|
Db 290 GPDLRRLVAR 300

RESULT 13
Q89RH2_BRAJA
ID Q89RH2_BRAJA PRELIMINARY; PRT; 580 AA.
AC Q89RH2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bll2800 protein.
GN OrderedLocusNames=bll2800;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,

RA Sasamoto S., Watanabe A., Idesawa K., Iriquchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC48065.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00481; PP2C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN 1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Transferase.
SQ SEQUENCE 580 AA; 64917 MW; 6AD3A06E6FAE143B CRC64;

Query Match 63.0%; Score 46; DB 2; Length 580;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
||| |||||
Db 355 IEGQTLRQWM 364

RESULT 14
O66272_9SPHN PRELIMINARY; PRT; 245 AA.
AC O66272;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14332;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB010981; BAA25791.1; -; Genomic_DNA.
DR HSSP; P02954; 1QOV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMS; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.

RA Sasamoto S., Watanabe A., Idesawa K., Iriquchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC48065.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00481; PP2C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN 1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Transferase.
SQ SEQUENCE 580 AA; 64917 MW; 6AD3A06E6FAE143B CRC64;

Query Match 63.0%; Score 46; DB 2; Length 580;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
||| |||||
Db 355 IEGQTLRQWM 364

RESULT 14
O66272_9SPHN PRELIMINARY; PRT; 245 AA.
AC O66272;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14332;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB010981; BAA25791.1; -; Genomic_DNA.
DR HSSP; P02954; 1QOV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMS; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.

FT NON TER 1
SQ SEQUENCE 245 AA; 27214 MW; 52B268713E199ABD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 245;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
||||| ||
Db 26 IEGPTLNPWL 35

RESULT 15
O82989_9SPHN PRELIMINARY; PRT; 249 AA.
ID O82989;
AC O82989;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3019;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB015708; BAA32995.1; -; Genomic_DNA.
DR HSSP; P02954; 1YST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMS; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON TER 1
SQ SEQUENCE 249 AA; 27702 MW; 4D68EDC82B7166AD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 249;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
||||| ||
Db 26 IEGPTLNPWL 35

RESULT 16
Q9XDV0_9SPHN PRELIMINARY; PRT; 278 AA.
ID Q9XDV0;
AC Q9XDV0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit.
GN Name=pufL;
OS Erythrobacter sp. MBIC3960.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=94771;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3960;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB027515; BAA78672.1; -; Genomic_DNA.
DR HSSP; P02954; 1YST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMS; TIGR01157; puFL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
DR SEQUENCE 278 AA; 30735 MW; 0BE618844B3C54FB CRC64;

Query Match 61.6%; Score 45; DB 2; Length 278;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
Db 55 IEGPTLNPWL 64

RESULT 17
Q7WLX1 BORPA
ID Q7WLX1 BORPA PRELIMINARY; PRT; 421 AA.
AC Q7WLX1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedLocusNames=BPP0223;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE39964.1; -; Genomic_DNA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
KW Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45579 MW; 13D6606AF1FDEC21 CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PTLRQWLAAR 13
Db 221 PSLRDWLAAR 230
|:|||||
|:|||||

RESULT 18
Q7WQU8 BORBR
ID Q7WQU8 BORBR PRELIMINARY; PRT; 421 AA.
AC Q7WQU8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedLocusNames=BB0227;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30725.1; -; Genomic_DNA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
KW Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45558 MW; A6CDBC98C731A49C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PTLRQWLAAR 13
Db 221 PSLRDWLAAR 230
|:|||||
|:|||||

RESULT 19
Q885P2 PSESM
ID Q885P2 PSESM PRELIMINARY; PRT; 756 AA.
AC Q885P2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dimethylsulfoxide reductase.
GN OrderedLocusNames=PSPT01789;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.

RL Science 300:1566-1569(2003).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,

RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,

RA Bera J., Kim M., Jin S., Fadrosch D., Vuong H., Overton II L.,

RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,

RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,

RA de Vazeilles A., White O., Salzberg S., Fraser C.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

RN [4]

RP NUCLEOTIDE SEQUENCE.

RA Buell R.;

RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE017067; AAP52546.1; -; Genomic_DNA.

DR EMBL; AC105932; AAX95563.1; -; Genomic_DNA.

DR Gramene; Q8LMK9; -;

DR InterPro; IPR005162; Retrotrans_gag.

DR Pfam; PF03732; Retrotrans_gag; 1.

KW Hypothetical protein; Polyprotein.

SQ SEQUENCE 760 AA; 82020 MW; C51F91AA2EB32A28 CRC64;

Query Match 60.3%; Score 44; DB 2; Length 760;

Best Local Similarity 46.2%; Pred. No. 1.1e+02;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13

Db 661 LHGPTLRQWMAVK 673

RESULT 26

KSGA_TREPA STANDARD; PRT; 285 AA.

AC O83357;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N',

DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High

DE level kasugamycin resistance protein ksgA) (Kasugamycin

DE dimethyltransferase).

GN Name=ksgA; OrderedLocusNames=TP0337;

OS Treponema pallidum.

OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=160;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Nichols;

RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,

RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,

RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,

RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,

RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,

RA Weidman J.F., Smith H.O., Venter J.C.;

RT "Complete genome sequence of Treponema pallidum, the syphilis

RT spirochete."

RL Science 281:375-388(1998).

CC -1- FUNCTION: Specifically dimethylates two adjacent adenosines in the

CC loop of a conserved hairpin near the 3' end of 16S rRNA in the 30S

CC particle. Its inactivation leads to kasugamycin resistance (By

CC similarity).

CC -1- SIMILARITY: Belongs to the rRNA adenine N-6-methyltransferase

CC family. KsgA subfamily.

CC

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AE001213; AAC65323.1; -; Genomic_DNA.

DR PIR; G71337; G71337.

DR TIGR; TP0337; -;

DR HAMAP; MF_00607; -; 1.

DR InterPro; IPR011530; ksgA.

DR InterPro; IPR001737; rRNA_meth_trans.

DR InterPro; IPR000051; SAM_Bd.

DR PANTHER; PTHR11727:SF6; ksgA; 1.

DR PANTHER; PTHR11727; rRNA_meth_trans; 1.

DR Pfam; PF00398; RrnaAD; 1.

DR SMART; SM00650; rADc; 1.

DR TIGRFAMs; TIGR00755; ksgA; 1.

DR PROSITE; PS01131; rRNA_A_DIMETH; 1.

KW Antibiotic resistance; Complete proteome; Methyltransferase;

KW rRNA processing; Transferase.

SQ SEQUENCE 285 AA; 32276 MW; 3AF0BCBE16EB5D4F CRC64;

Query Match 58.9%; Score 43; DB 1; Length 285;

Best Local Similarity 64.3%; Pred. NO. 59;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14

Db 98 IEGDVLQQWHAARA 111

RESULT 27

Q7UQE4_RHOBA PRELIMINARY; PRT; 297 AA.

AC Q7UQE4;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=RB6375;

OS Rhodopirellula baltica.

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;

OC Planctomycetaceae; Pirellula.

OX NCBI_TaxID=117;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=1;

RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;

RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,

RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,

RA Schlesner H., Amann R., Reinhardt R.;

RT "Complete genome sequence of the marine planctomycete Pirellula sp.

RT strain 1."

RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).

DR EMBL; BX294144; CAD74759.1; -; Genomic_DNA.

DR InterPro; IPR000194; ATPase_a/bcentre.

DR InterPro; IPR003169; GYF.

DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.

DR PROSITE; PS50829; GYF; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 297 AA; 31805 MW; 475F670F02C78E9B CRC64;

Query Match 58.9%; Score 43; DB 2; Length 297;

Best Local Similarity 50.0%; Pred. No. 62;

Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGPTLRQWLA 11

Db 176 DGPTMKQWIS 185

RESULT 28

Q8EJ00_SHEON PRELIMINARY; PRT; 306 AA.

ID Q8EJ00_SHEON


```
AC Q8EJ00;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Prophage MuSol, major head subunit, putative.
GN OrderedLocusNames=SO0675;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AB015513; AAN53753.1; -; Genomic_DNA.
DR TIGR; SO0675; -.
KW Complete proteome.
SQ SEQUENCE 306 AA; 34370 MW; F54CCA118AA288CB CRC64;

Query Match 58.9%; Score 43; DB 2; Length 306;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 PTLRQWLAAR 13
||:|:|:|
Db 54 PTMREWIGAR 63

RESULT 29
Q8ZYT5 PYRAE
ID Q8ZYT5 PYRAE PRELIMINARY; PRT; 354 AA.
AC Q8ZYT5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0634.
GN OrderedLocusNames=PAE0634;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009776; AAL62908.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 354 AA; 38642 MW; C5799F975B972941 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 354;
Best Local Similarity 61.5%; Pred. No. 74;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
|:|:|:|:|
Db 84 IDRPGLEQWLASR 96

RESULT 29
Q8ZYT5 PYRAE
ID Q8ZYT5 PYRAE PRELIMINARY; PRT; 354 AA.
AC Q8ZYT5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0634.
GN OrderedLocusNames=PAE0634;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009776; AAL62908.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 354 AA; 38642 MW; C5799F975B972941 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 354;
Best Local Similarity 61.5%; Pred. No. 74;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
|:|:|:|:|
Db 84 IDRPGLEQWLASR 96
```

```
RESULT 30
Q82PX5 STRAW
ID Q82PX5 STRAW PRELIMINARY; PRT; 377 AA.
AC Q82PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV747;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.2114331198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC68457.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 377 AA; 41307 MW; 0253176AAAE62F3 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 377;
Best Local Similarity 61.5%; Pred. No. 79;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
|:|:|:|:|
Db 168 MEGPDLRAWLPNR 180

RESULT 31
O83436 TREPA
ID O83436 TREPA PRELIMINARY; PRT; 683 AA.
AC O83436;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein TP0421.
GN OrderedLocusNames=TP0421;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artiach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
```

RL Science 281:375-388(1998).
DR EMBL; AB001220; AAC65409.1; -; Genomic_DNA.
DR PIR; B71325; B71325.
DR TIGR; TP0421; -.
DR GO; GO:0005488; F:binding; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR013017; NHL_rep.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF01436; NHL; 5.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
DR Complete proteome; Repeat; TPR repeat.
KW SEQUENCE 683 AA; 74519 MW; F91407FA7094AAD1 CRC64;
SQ

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 1.4e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEPTLRQWLAAR 13
||| | |||
Db 89 IEGAALHQWGAAR 101

RESULT 32
Q617Q3 CAEBR
ID Q617Q3 CAEBR PRELIMINARY; PRT; 728 AA.
AC Q617Q3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14874.
GN Name=CBG14874;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; CAAC01000068; CAE68902.1; -; Genomic_DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Complete proteome; Hypothetical protein; Metal-binding;
KW Ub1 conjugation pathway; Zinc; Zinc-finger.
KW Hypothetical protein; Metal-binding; Ub1 conjugation pathway; Zinc;
KW Zinc-finger.
SQ SEQUENCE 728 AA; 82017 MW; 97016C238A26F96B CRC64;

Query Match 58.9%; Score 43; DB 2; Length 728;
Best Local Similarity 63.6%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
||| | |||
Db 614 GPCLRKWLAVK 624

RESULT 33
Q95Y82 CAEEL
ID Q95Y82 CAEEL PRELIMINARY; PRT; 754 AA.
AC Q95Y82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Y119C1B.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AC006712; AAK39324.1; -; Genomic_DNA.
DR Ensembl; Y119C1B.5; Caenorhabditis elegans.
DR WormBase; WBGene00022471; Y119C1B.5.
DR WormPep; Y119C1B.5; CE27234.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Complete proteome; Hypothetical protein; Metal-binding;
KW Ub1 conjugation pathway; Zinc; Zinc-finger.
SQ SEQUENCE 754 AA; 85324 MW; 41BAA9297FA3BF05 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 754;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAR 13
||| | |||
Db 626 GPCLRKWLAVK 636

RESULT 34
Q527F0 MAGGR
ID Q527F0 MAGGR PRELIMINARY; PRT; 1157 AA.
AC Q527F0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG05740.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,

RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihaliev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=70-15;

RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=70-15;

RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AACU01000566; EAA54949.1; -; Genomic_DNA.

KW Hypothetical protein.

SQ SEQUENCE 1157 AA; 124735 MW; 2EECD2BCB30618B CRC64;

Query Match 58.9%; Score 43; DB 2; Length 1157;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAA 12

Db 150 VDGPKLKEWLQA 161

RESULT 35

Q8Y0I5 RALSO

ID Q8Y0I5_RALSO PRELIMINARY; PRT; 91 AA.

AC Q8Y0I5;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical protein Rsc1059.

GN OrderedLocusNames=RS04149;

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Burkholderiaceae; Ralstonia.

OX NCBI_TaxID=305;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GMI1000;

RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;

RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,

RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,

RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,

RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,

RA Weissenbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";

RL Nature 415:497-502(2002).

DR EMBL; AL646062; CAD14761.1; -; Genomic DNA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR008554; GLRX2.

DR InterPro; IPR012336; Thioredoxin-like.

DR InterPro; IPR012335; Thioredoxin_fold.

DR Pfam; PF05768; DUF836; 1.

KW Complete proteome.

SQ SEQUENCE 91 AA; 10321 MW; 2B4DFFEB37A528AD CRC64;

Query Match 57.5%; Score 42; DB 2; Length 91;

Best Local Similarity 46.2%; Pred. No. 28;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAA 13

Db 75 LDGPAVQAWLAAQ 87

RESULT 36

Q8N9N4 HUMAN

ID Q8N9N4_HUMAN PRELIMINARY; PRT; 126 AA.

AC Q8N9N4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypothetical protein FLJ36840.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX PubMed=14702039; DOI=10.1038/ngl285;

RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,

RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,

RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,

RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,

RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,

RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,

RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,

RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,

RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,

RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,

RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,

RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,

RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,

RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,

RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,

RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,

RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,

RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,

RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,

RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,

RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,

RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,

RT "Complete sequencing and characterization of 21,243 full-length human

RT cDNAs.";

RL Nat. Genet. 36:40-45(2004).

DR EMBL; AK094159; BAC04297.1; -; mRNA.

DR Ensembl; ENSG00000188274; Homo sapiens.

SQ SEQUENCE 126 AA; 14003 MW; AF10E9375A3D9C7E CRC64;

Query Match 57.5%; Score 42; DB 2; Length 126;
Best Local Similarity 58.3%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAARA 14
Db 108 GPDLLRWAGSRA 119

RESULT 37
Q4SMZ4 TETNG
ID Q4SMZ4_TETNG PRELIMINARY; PRT; 127 AA.
AC Q4SMZ4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 6 SCAF14544, whole genome shotgun sequence.
GN ORFNames=GSTENG00015515001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014544; CAF97988.1; -; Genomic_DNA.
SQ SEQUENCE 127 AA; 14785 MW; 0DE1B75D22B88BC8 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 127;
Best Local Similarity 57.1%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
Db 59 IEDPALTRWVHARA 72

RESULT 38
Q8XPQ9 RALSO
ID Q8XPQ9_RALSO PRELIMINARY; PRT; 252 AA.
AC Q8XPQ9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
GN OrderedLocusNames=RSp1579; ORFNames=RS02135;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.

OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- SIMILARITY: Contains 1 HTH luxR-type DNA-binding domain.
DR EMBL; AL646085; CAD18730.1; -; Genomic_DNA.
DR HSSP; P11470; 1FSE.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00196; GerE; 1.
DR PRINTS; PR01590; HTHFIS.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00421; HTH_LUXR; 1.
KW Complete proteome; DNA-binding; Plasmid; Transcription;
KW Transcription regulation.
SQ SEQUENCE 252 AA; 27666 MW; 483403EE326F7C2E CRC64;

Query Match 57.5%; Score 42; DB 2; Length 252;
Best Local Similarity 53.8%; Pred. No. 78;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
Db 76 IDTPIMRRWLATR 88

RESULT 39
P90433 SIVCZ
ID P90433_SIVCZ PRELIMINARY; PRT; 313 AA.
AC P90433;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Truncated reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Smith J.M., Kraiselburd E.N., Torres J.V.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U83413; AAB41428.1; -; Genomic_DNA.
DR HSSP; Q07387; 1TCW.
DR SMR; P90433; 69-167, 173-313.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.

DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase.
FT NON_TER 1
SQ SEQUENCE 313 AA; 34675 MW; 5A0BB016783FC8A6 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 313;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

RESULT 40
Q855N9_9CAUD
ID Q855N9_9CAUD PRELIMINARY; PRT; 325 AA.
AC Q855N9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp74.
OS Mycobacteriophage Che9d.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205876;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes."
RL Cell 113:171-182(2003).
DR EMBL; AY129336; AAN07992.1; -; Genomic_DNA.
SQ SEQUENCE 325 AA; 35999 MW; 04265796D0B4FC1D CRC64;

Query Match 57.5%; Score 42; DB 2; Length 325;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAARA 14
Db 302 DGPTVQEALARA 314

RESULT 41
P95613_RHIGA
ID P95613_RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAMBI;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstroem K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.
DR EMBL; Y08963; CAA70157.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LysR; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 57.5%; Score 42; DB 2; Length 326;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQWLAAR 13
Db 204 KGPSLEQWLSSQ 215

RESULT 42
Q7FAN4_ORYSA
ID Q7FAN4_ORYSA PRELIMINARY; PRT; 375 AA.
AC Q7FAN4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE OSJNBb0060E08.2 protein.
GN Name=OSJNBb0060E08.2;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4."
RL Nature 420:316-320(2002).
DR EMBL; AL606669; CAE04739.1; -; Genomic_DNA.
DR Gramene; Q7XPP6; -.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR002083; MATH.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00917; MATH; 1.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00061; MATH; 1.
DR PROSITE; PS50097; BTB; 1.
DR PROSITE; PS50144; MATH; 1.
SQ SEQUENCE 375 AA; 41043 MW; 20FC6E99E4750816 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
Db 138 MERPRLRQWLLRR 150

RESULT 43
Q7XPP6_ORYSA

KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
KW Nuclease; Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferase; Zinc; Zinc-finger.
FT CHAIN 1 167
FT DOMAIN 88 157 Peptidase A2.
FT DOMAIN 211 401 Reverse transcriptase.
FT DOMAIN 600 723 RNase H.
FT DOMAIN 779 930 Integrase catalytic.
FT ZN_FING 729 770 Integrase-type.
FT DNA_BIND 949 996 Integrase-type.
FT ACT_SITE 93 93 By similarity.
SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match 57.5%; Score 42; DB 1; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
||| ||||
Db 184 EGPQLRQW 191

RESULT 48
P89154_SIVCZ
ID P89154_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC P89154;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Pol protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SIVsmE543;
RX MEDLINE=97151152; PubMed=8995688;
RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,
RA Elkins W.R., Montefiori D.C.;
RT "A molecularly cloned, pathogenic, neutralization-resistant simian
RT immunodeficiency virus, SIVsmE543-3.";
RL J. Virol. 71:1608-1620(1997).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SIVsmE543;
RA Ourmanov I.K., Deghani H., Kuwata T., Hirsch V.M.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).

CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.

CC EMBL; U72748; AAC56559.2; -; Genomic_DNA.
CC SMR; P89154; 69-167, 781-939.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.

DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1 1
SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
||| ||||
Db 184 EGPQLRQW 191

RESULT 49

Q7ZBR7_SIVCZ
ID Q7ZBR7_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Deghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).

CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.

CC EMBL; AY221514; AA067307.1; -; Genomic_DNA.
CC HSSP; P04584; 1MU2.
CC SMR; Q7ZBR7; 69-167, 781-939.
CC MEROPS; A02.002; -.
CC GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF06815; RVT_thumb; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS0141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS1027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115341 MW; A886525DFFIBE26F CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

RESULT 50
ID Q7ZBR5_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian

RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221515; AAC67309.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

Search completed: May 12, 2006, 10:52:00
Job time : 114.077 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 65.3846 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	92.9	11	2	AAY22365 TPO recep
2	26	92.9	11	2	AAY22370 TPO recep
3	26	92.9	11	3	AAB16986 TPO-mimet
4	26	92.9	11	3	AAB16988 TPO-mimet
5	26	92.9	11	5	ABB72872 TPO mimet
6	26	92.9	11	5	ABB72874 TPO mimet
7	26	92.9	11	7	ADJ73023 TPO mimet
8	26	92.9	11	7	ADJ73025 TPO mimet
9	26	92.9	11	8	ADJ52658 CH1 delet
10	26	92.9	11	8	ADJ52660 CH1 delet
11	26	92.9	11	8	ADJ51619 CH1 delet
12	26	92.9	11	8	ADJ51621 CH1 delet
13	26	92.9	13	2	AAY22407 TPO recep
14	26	92.9	95	3	AAY65258 Human 5'
15	26	92.9	95	8	ADU72822 Signal pe
16	26	92.9	95	9	ADZ73813 Human inc
17	26	92.9	357	6	ABU48350 Protein e
18	26	92.9	465	8	ADI67225 Lactobaci
19	26	92.9	486	4	ABB71566 Drosophil
20	26	92.9	510	7	ABO66441 Klebsiell
21	26	92.9	2559	9	ABM96913 M. xanthu
22	25	89.3	9	3	AAB16977 TPO-mimet
23	25	89.3	9	5	ABB72863 TPO mimet
24	25	89.3	9	6	ABG71745 TPO recep

98	24	85.7	264	7	ADM04596	Adm04596 Human pro	171	24	85.7	655	8	ADJ49379	Oil-assoc
99	24	85.7	266	3	AAG34624	Aag34624 Arabidops	172	24	85.7	664	7	ADJ69530	Human hea
100	24	85.7	267	7	ABO63689	Abo63689 Klebsiell	173	24	85.7	666	7	ABO83796	Pseudomon
101	24	85.7	292	4	AAB80103	Aab80103 Corynebac	174	24	85.7	698	8	ADI19358	Soybean t
102	24	85.7	304	6	ABU16626	Abu16626 Protein e	175	24	85.7	698	8	ADJ65943	soybean t
103	24	85.7	315	3	AAG26425	Aag26425 Arabidops	176	24	85.7	710	6	ABR58651	Human can
104	24	85.7	315	3	AAG46503	Aag46503 Arabidops	177	24	85.7	723	6	ABU24903	Protein e
105	24	85.7	316	5	ADI16829	Adi16829 Murine NO	178	24	85.7	737	8	ADR10261	Human pro
106	24	85.7	316	7	ABW02143	Abw02143 Human GPC	179	24	85.7	741	6	ABU21306	Protein e
107	24	85.7	317	2	AAR45363	Aar45363 Naei rest	180	24	85.7	802	6	ADA54442	Human pro
108	24	85.7	324	5	ABP56427	Abp56427 Human glu	181	24	85.7	874	4	ABG26108	Novel hum
109	24	85.7	342	9	AEB41377	Aeb41377 L. pneumo	182	24	85.7	891	3	AAB26596	Synechocy
110	24	85.7	345	3	AAG26424	Aag26424 Arabidops	183	24	85.7	897	6	ADB11560	Alloiococ
111	24	85.7	345	3	AAG46502	Aag46502 Arabidops	184	24	85.7	924	8	ADP99020	C. albica
112	24	85.7	350	4	AAB50667	Aab50667 C. elegan	185	24	85.7	966	6	ABM70145	Photorhab
113	24	85.7	351	4	AAB50666	Aab50666 C. elegan	186	24	85.7	993	4	ABB65429	Drosophil
114	24	85.7	351	9	AEB38085	Aeb38085 L. pneumo	187	24	85.7	996	9	AEB21645	Adhesin-1
115	24	85.7	360	4	AAG91578	Aag91578 C glutami	188	24	85.7	997	4	ABG15941	Novel hum
116	24	85.7	363	3	AAG26423	Aag26423 Arabidops	189	24	85.7	1001	8	ADI28617	Mycoplasm
117	24	85.7	363	3	AAG46501	Aag46501 Arabidops	190	24	85.7	1166	8	ADN20571	Bacterial
118	24	85.7	363	5	ABB93830	Abb93830 Herbicida	191	24	85.7	2951	6	ABM67717	Photorhab
119	24	85.7	363	8	ADJ49252	Adj49252 Oil-assoc	192	24	85.7	2975	6	ABM68759	Photorhab
120	24	85.7	363	8	ADJ49645	Adj49645 Oil-assoc	193	24	85.7	3041	6	ABM70325	Photorhab
121	24	85.7	363	8	ADJ49605	Adj49605 Oil-assoc	194	24	85.7	4560	6	ABM67454	Photorhab
122	24	85.7	385	3	AAG52926	Aag52926 Arabidops	195	23	82.1	10	2	AAY22355	TPO recep
123	24	85.7	385	3	AAG41437	Aag41437 Arabidops	196	23	82.1	10	3	AAB16979	TPO-mimet
124	24	85.7	430	6	ADB08462	Adb08462 Alloiococ	197	23	82.1	10	5	ABB72865	TPO mimet
125	24	85.7	430	8	ADJ27093	Adj27093 Alloiococ	198	23	82.1	10	7	ADJ73016	TPO mimet
126	24	85.7	432	4	AAU51431	Aau51431 Propionib	199	23	82.1	10	8	ADJ52651	CH1 delet
127	24	85.7	432	6	ABM47950	Abm47950 Propionib	200	23	82.1	10	8	ADJ51612	CH1 delet
128	24	85.7	435	3	AAB10917	Aab10917 Human reg	201	23	82.1	12	5	ABB05304	Polyureth
129	24	85.7	441	7	ADM05130	Adm05130 Human pro	202	23	82.1	15	2	AAY22395	TPO recep
130	24	85.7	456	6	ADB11562	Adb11562 Alloiococ	203	23	82.1	15	2	AAY22403	TPO recep
131	24	85.7	461	3	AAG11836	Aag11836 Arabidops	204	23	82.1	15	2	AAY22373	TPO recep
132	24	85.7	461	3	AAG46490	Aag46490 Arabidops	205	23	82.1	15	2	AAY22390	TPO recep
133	24	85.7	462	3	AAG46489	Aag46489 Arabidops	206	23	82.1	15	2	AAY22378	TPO recep
134	24	85.7	462	3	AAG11835	Aag11835 Arabidops	207	23	82.1	15	2	AAY22397	TPO recep
135	24	85.7	464	8	ADX94959	Adx94959 Plant ful	208	23	82.1	25	4	AAM16749	Peptide #
136	24	85.7	465	3	AAG46491	Aag46491 Arabidops	209	23	82.1	25	4	ABB35735	Peptide #
137	24	85.7	465	3	AAG11834	Aag11834 Arabidops	210	23	82.1	25	4	AAM29237	Peptide #
138	24	85.7	465	8	ADT56526	Adt56526 Plant pol	211	23	82.1	25	4	ABB30570	Peptide #
139	24	85.7	468	6	ABJ25769	Abj25769 Aspergill	212	23	82.1	25	4	ABB21161	Protein #
140	24	85.7	469	6	ABJ26369	Abj26369 Aspergill	213	23	82.1	25	4	AAM68926	Human bon
141	24	85.7	471	3	AAB21094	Aab21094 Yeast HEL	214	23	82.1	25	4	AAM56544	Human bra
142	24	85.7	471	6	ABR52624	Abr52624 Protein s	215	23	82.1	25	4	ABG50591	Human liv
143	24	85.7	471	7	ADB84329	Adb84329 Yeast rep	216	23	82.1	25	4	AAM04466	Peptide #
144	24	85.7	471	7	ADK62264	Adk62264 Disease t	217	23	82.1	25	5	ABG38507	Human pep
145	24	85.7	481	4	ABB71798	Abb71798 Drosophil	218	23	82.1	62	5	ABP32688	Human ORF
146	24	85.7	481	4	AAE03541	Aae03541 Drosophil	219	23	82.1	62	5	ADG79445	Human sec
147	24	85.7	481	7	ADB04328	Adb04328 Drosophil	220	23	82.1	62	5	ADG79538	Human sec
148	24	85.7	498	5	ABP73641	Abp73641 Candida a	221	23	82.1	68	8	ABO57882	Human gen
149	24	85.7	505	6	ABR53169	Abr53169 Protein s	222	23	82.1	76	4	AAU66802	Propionib
150	24	85.7	505	7	ADK63188	Adk63188 Disease t	223	23	82.1	76	6	ABM63321	Propionib
151	24	85.7	520	7	ABO80725	Abo80725 Pseudomon	224	23	82.1	79	9	ABM92162	M. xanthu
152	24	85.7	527	8	ADS23096	Ads23096 Bacterial	225	23	82.1	109	3	AAG01169	Human sec
153	24	85.7	530	5	ABP69308	Abp69308 Human pol	226	23	82.1	111	8	ADI57234	FEN1 domi
154	24	85.7	543	8	ADV89911	Adv89911 Streptoco	227	23	82.1	114	3	AAB40430	Human ORF
155	24	85.7	543	8	ADV83301	Adv83301 Streptoco	228	23	82.1	114	5	ABP09147	Human ORF
156	24	85.7	543	8	ADV81164	Adv81164 Streptoco	229	23	82.1	118	4	ABG11928	Novel hum
157	24	85.7	543	8	ADX67350	Adx67350 Plant ful	230	23	82.1	136	3	AA996838	pCpB-DQVD
158	24	85.7	551	2	AAW18790	Aaw18790 Corrected	231	23	82.1	142	6	ADB08396	Alloiococ
159	24	85.7	572	8	ADX74519	Adx74519 Plant ful	232	23	82.1	147	7	ABO66415	Klebsiell
160	24	85.7	585	4	ABB60953	Abb60953 Drosophil	233	23	82.1	153	4	AAO05794	Human pol
161	24	85.7	591	4	ABB63312	Abb63312 Drosophil	234	23	82.1	154	8	ADN41234	Human sec
162	24	85.7	600	8	ADS29505	Ads29505 Bacterial	235	23	82.1	154	8	ADA07510	Novel hum
163	24	85.7	600	9	ABM96218	Abm96218 M. xanthu	236	23	82.1	164	3	AAB44991	Human sec
164	24	85.7	624	5	ABP65819	Abp65819 Bifidobac	237	23	82.1	164	3	AAB44992	Human sec
165	24	85.7	628	8	ADN20030	Adn20030 Bacterial	238	23	82.1	171	6	ABU26770	Protein e
166	24	85.7	645	7	ADC10136	Adc10136 Human NOV	239	23	82.1	172	2	AA935324	Chlamydia
167	24	85.7	654	6	ABR53790	Abr53790 Protein s	240	23	82.1	177	6	ABU27123	Protein e
168	24	85.7	654	7	ADK62018	Adk62018 Disease t	241	23	82.1	182	6	ADA36082	Acinetoba
169	24	85.7	654	8	ADJ50308	Adj50308 Oil-assoc	242	23	82.1	192	8	ADX75919	Plant ful
170	24	85.7	654	8	ADS43755	Ads43755 Bacterial	243	23	82.1	206	3	AAG07266	Arabidops

Adj49379	Oil-assoc
Adj69530	Human hea
Abo83796	Pseudomon
Adl19358	Soybean t
Adj65943	soybean t
Abr58651	Human can
Abu24903	Protein e
Adr10261	Human pro
Abu21306	Protein e
Ada54442	Human pro
Abg26108	Novel hum
Aab26596	Synechocy
Adb11560	Alloiococ
Adp99020	C. albica
Abm70145	Photorhab
Abb65429	Drosophil
Aeb21645	Adhesin-1
Abg15941	Novel hum
Adi28617	Mycoplasm
Adn20571	Bacterial
Abm67717	Photorhab
Abm68759	Photorhab
Abm70325	Photorhab
Abm67454	Photorhab
Aay22355	TPO recep
Aab16979	TPO-mimet
Abb72865	TPO mimet
Adj73016	TPO mimet
Adj52651	CH1 delet
Adj51612	CH1 delet
Abb05304	Polyureth
Aay22395	TPO recep
Aay22403	TPO recep
Aay22373	TPO recep
Aay22390	TPO recep
Aay22378	TPO recep
Aay22397	TPO recep
Aam16749	Peptide #
Abb35735	Peptide #
Aam29237	Peptide #
Abb30570	Peptide #
Abb21161	Protein #
Aam68926	Human bon
Aam56544	Human bra
Abg50591	Human liv
Aam04466	Peptide #
Abg38507	Human pep
Abp32688	Human ORF
Adg79445	Human sec
Adg79538	Human sec
Abo57882	Human gen
Aau66802	Propionib
Abm63321	Propionib
Abm92162	M. xanthu
Aag01169	Human sec
Adi57234	FEN1 domi
Aab40430	Human ORF
Abp09147	Human ORF
Abg11928	Novel hum
AA996838	pCpB-DQVD
Adb08396	Alloiococ
ABO66415	Klebsiell
Aao05794	Human pol
ADN41234	Human sec
ADA07510	Novel hum
AAB44991	Human sec
AAB44992	Human sec
ABU26770	Protein e
Aay35324	Chlamydia
Abu27123	Protein e
Ada36082	Acinetoba
Adx75919	Plant ful
Aag07266	Arabidops

244	23	82.1	212	9	AEB40014	Aeb40014 L. pneumo	317	23	82.1	403	8	ADQ89864	Adq89864 Antagonis
245	23	82.1	214	7	ADC96874	Adc96874 E. faeciu	318	23	82.1	403	8	ADS10466	Ads10466 Human the
246	23	82.1	215	9	AEb36622	Aeb36622 L. pneumo	319	23	82.1	403	9	ADZ14402	Adz14402 Mouse squ
247	23	82.1	218	6	ABU35228	Abu35228 Protein e	320	23	82.1	405	9	ABM95316	Abm95316 M. xanthu
248	23	82.1	220	9	AEb37357	Aeb37357 L. pneumo	321	23	82.1	408	4	AAB48304	Aab48304 Human ZF2
249	23	82.1	220	9	AEb40677	Aeb40677 L. pneumo	322	23	82.1	409	2	AAR25693	Aar25693 Heat resi
250	23	82.1	220	9	AEb40670	Aeb40670 L. pneumo	323	23	82.1	409	2	AAR25693	Aar25693 Heat resi
251	23	82.1	224	9	AEb37350	Aeb37350 L. pneumo	324	23	82.1	410	8	AAW03544	Aaw03544 Bacillus
252	23	82.1	226	4	ABG02211	Abg02211 Novel hum	325	23	82.1	411	4	ADR10146	Adr10146 Human pro
253	23	82.1	237	3	AAG36022	Aag36022 Zea mays	326	23	82.1	414	5	ABG13838	Abg13838 Novel hum
254	23	82.1	245	3	AAG07265	Aag07265 Arabidops	327	23	82.1	417	7	ABG70045	Abg70045 Human pre
255	23	82.1	247	3	AAy37499	Aay37499 Protein w	328	23	82.1	418	4	ABO73714	Abo73714 Pseudomon
256	23	82.1	247	3	AAB42519	Aab42519 Human ORF	329	23	82.1	419	4	ADS21594	Ads21594 Bacterial
257	23	82.1	254	3	AAG20958	Aag20958 Arabidops	330	23	82.1	420	4	ABG13105	Abg13105 Novel hum
258	23	82.1	265	5	ABJ04666	Abj04666 Mycobacte	331	23	82.1	420	8	AAM39925	Aam39925 Human pol
259	23	82.1	281	3	AAG36021	Aag36021 Zea mays	332	23	82.1	423	8	ADS24829	Ads24829 Bacterial
260	23	82.1	291	3	AAG36020	Aag36020 Zea mays	333	23	82.1	425	4	ADT57937	Adt57937 Plant pol
261	23	82.1	303	2	AAW06422	Aaw06422 Yeast sn-	334	23	82.1	426	8	AAM41711	Aam41711 Human pol
262	23	82.1	303	2	AAW06421	Aaw06421 Yeast sn-	335	23	82.1	426	8	ADJ49912	Adj49912 Oil-assoc
263	23	82.1	303	2	AAW88102	Aaw88102 Transfera	336	23	82.1	426	8	ADJ50126	Adj50126 Oil-assoc
264	23	82.1	303	3	AAy99490	Aay99490 Yeast acy	337	23	82.1	428	8	ADS27875	Ads27875 Bacterial
265	23	82.1	303	5	AAE15286	Aae15286 Yeast LPA	338	23	82.1	432	8	ADJ49280	Adj49280 Oil-assoc
266	23	82.1	303	6	ABR52952	Abr52952 Protein s	339	23	82.1	432	8	ADN26876	Adn26876 Bacterial
267	23	82.1	303	6	ABU24010	Abu24010 Protein e	340	23	82.1	433	3	ADN26675	Adn26675 Bacterial
268	23	82.1	303	7	ADK62618	Adk62618 Disease t	341	23	82.1	436	6	AAG20957	Aag20957 Arabidops
269	23	82.1	303	8	ADJ93787	Adj93787 Yeast lys	342	23	82.1	439	4	ABU41319	Abu41319 Protein e
270	23	82.1	303	8	ADN49471	Adn49471 Yeast lys	343	23	82.1	439	8	AAB93788	Aab93788 Human pro
271	23	82.1	303	8	ADO04610	Ado04610 Yeast lys	344	23	82.1	446	8	ADY23217	Ady23217 Plant ful
272	23	82.1	303	8	ADN18936	Adn18936 Bacterial	345	23	82.1	448	4	ADS28594	Ads28594 Bacterial
273	23	82.1	308	3	AAB43188	Aab43188 Human ORF	346	23	82.1	448	4	AAG82323	Aag82323 S. epider
274	23	82.1	315	5	ADI16827	Adi16827 Murine NO	347	23	82.1	460	4	AAM78655	Aam78655 Human GTP
275	23	82.1	315	7	ABW02144	Abw02144 Human GPC	348	23	82.1	460	4	AAB68506	Aab68506 Human GTP
276	23	82.1	315	8	ADM42847	Adm42847 Murine od	349	23	82.1	468	5	ABP43604	Abp43604 Clone MGC
277	23	82.1	319	4	ABB68087	Abb68087 Drosophil	350	23	82.1	469	8	ADT59033	Adt59033 Plant pol
278	23	82.1	319	9	AEb41680	Aeb41680 L. pneumo	351	23	82.1	470	6	AAM79639	Aam79639 Human pro
279	23	82.1	321	4	ABB61436	Abb61436 Drosophil	352	23	82.1	470	6	ABU34745	Abu34745 Protein e
280	23	82.1	321	4	AAB64718	Aab64718 Human sec	353	23	82.1	471	5	ADH22542	Adh22542 Human tra
281	23	82.1	321	4	AAB64717	Aab64717 Gene 5 hu	354	23	82.1	471	4	AAY90287	Aay90287 Human pep
282	23	82.1	321	4	AAB65978	Aab65978 Human sec	355	23	82.1	471	5	AAB93481	Aab93481 Human pro
283	23	82.1	321	4	AAE65977	Aae65977 Human sec	356	23	82.1	474	9	AAM47764	Aam47764 F-cassett
284	23	82.1	328	9	AEb38409	Aeb38409 L. pneumo	357	23	82.1	481	9	AEB39197	Aeb39197 L. pneumo
285	23	82.1	337	9	AEA99848	Aea99848 L. helvet	358	23	82.1	488	7	AEb35747	Aeb35747 L. pneumo
286	23	82.1	344	4	AAG72619	Aag72619 Murine OR	359	23	82.1	491	6	ABO66647	Abo66647 Klebsiell
287	23	82.1	345	3	AAB42984	Aab42984 Human ORF	360	23	82.1	502	6	ABU26104	Abu26104 Protein e
288	23	82.1	349	9	ADW26685	Adw26685 Fructo-ol	361	23	82.1	503	2	ABU26099	Abu26099 Protein e
289	23	82.1	350	4	AAB94228	Aab94228 Human pro	362	23	82.1	503	6	AAR72508	Aar72508 Rat chole
290	23	82.1	350	4	ABG08347	Abg08347 Novel hum	363	23	82.1	503	7	ABP57062	Abp57062 Rat CYP7A
291	23	82.1	353	8	ADQ67803	Adq67803 Novel hum	364	23	82.1	504	2	Add48630	Add48630 Rat Prote
292	23	82.1	355	5	ABP66105	Abp66105 Bifidobac	365	23	82.1	504	2	Adi19746	Adi19746 Rat CYP7A
293	23	82.1	357	5	ABP63001	Abp63001 Human pol	366	23	82.1	504	8	AAR72509	Aar72509 Hamster c
294	23	82.1	360	5	ABP41132	Abp41132 Human ova	367	23	82.1	515	4	Aau33948	Aau33948 Staphyloc
295	23	82.1	368	4	ABG28311	Abg28311 Novel hum	368	23	82.1	523	4	Abb65104	Abb65104 Drosophil
296	23	82.1	370	4	AAB94432	Aab94432 Human pro	369	23	82.1	525	7	Ade71266	Ade71266 Novel hum
297	23	82.1	378	7	ADK66908	Adk66908 Human toc	370	23	82.1	534	6	Aae30009	Aae30009 Tobacco c
298	23	82.1	383	2	AAW42447	Aaw42447 C. hetero	371	23	82.1	537	6	Aae30008	Aae30008 Tobacco c
299	23	82.1	385	4	ABB67960	Abb67960 Drosophil	372	23	82.1	544	6	Abu35831	Abu35831 Protein e
300	23	82.1	389	4	AAU34818	Aau34818 E. coli c	373	23	82.1	544	8	Adm97949	Adm97949 GFTpsD1.
301	23	82.1	389	8	ADJ64184	Adj64184 Escherich	374	23	82.1	548	9	Ady58795	Ady58795 GFTpsD2.
302	23	82.1	389	8	ADK13790	Adk13790 E. coli i	375	23	82.1	555	8	Adm97948	Adm97948 GFTpsB. 7
303	23	82.1	397	4	ABG30029	Abg30029 Novel hum	376	23	82.1	557	6	Abu15931	Abu15931 Protein e
304	23	82.1	400	5	ADI16915	Adi16915 Rat NOVX	377	23	82.1	557	9	Adw94871	Adw94871 Prolifera
305	23	82.1	403	4	AAG67025	Aag67025 Human SPF	378	23	82.1	566	1	Aap82595	Aap82595 Amylase f
306	23	82.1	403	4	AAG67024	Aag67024 Rat SPF.	379	23	82.1	566	2	Aar08263	Aar08263 Amylase f
307	23	82.1	403	5	ABG61931	Abg61931 Prostata	380	23	82.1	568	6	Abu43665	Abu43665 Protein e
308	23	82.1	403	5	ABP41422	Abp41422 Human ova	381	23	82.1	571	6	Adb06954	Adb06954 Alloiococ
309	23	82.1	403	5	ADI16913	Adi16913 Rat NOVX	382	23	82.1	572	5	Abp43025	Abp43025 Protein e
310	23	82.1	403	5	ADI16912	Adi16912 Human NOV	383	23	82.1	573	6	Abp40097	Abp40097 Staphyloc
311	23	82.1	403	5	ADI16914	Adi16914 Murine NO	384	23	82.1	573	8	Ads05014	Ads05014 Staphyloc
312	23	82.1	403	7	ADK66910	Adk66910 Human toc	385	23	82.1	574	4	Aau37358	Aau37358 Staphyloc
313	23	82.1	403	7	ABO66445	Abo66445 Klebsiell	386	23	82.1	574	4	Aau36733	Aau36733 Staphyloc
314	23	82.1	403	8	ADJ75589	Adj75589 Marker ge	387	23	82.1	574	6	Abj18989	Abj18989 Pathogen
315	23	82.1	403	8	ADN02622	Adn02622 Liver dis	388	23	82.1	574	6	Abm73119	Abm73119 Staphyloc
316	23	82.1	403	8	ADQ19056	Adq19056 Human sof	389	23	82.1	584	7	Adc94937	Adc94937 E. faeciu

390	23	82.1	610	8	ADN20025	Adn20025 Bacterial	463	23	82.1	896	8	ADQ89162	Adq89162 Human uro
391	23	82.1	615	7	ADF07764	Adf07764 Bacterial	464	23	82.1	896	8	ADR44918	Adr44918 Polypepti
392	23	82.1	617	7	ABO67804	Abo67804 Klebsiell	465	23	82.1	900	3	AAG47642	Aag47642 Arabidops
393	23	82.1	619	8	ADN27232	Adn27232 Bacterial	466	23	82.1	908	4	ABB65159	Abb65159 Drosophil
394	23	82.1	622	6	ADB06956	Adb06956 Alloiococ	467	23	82.1	916	7	ADI21204	Adi21204 Novel hum
395	23	82.1	633	4	ABG28306	Abg28306 Novel hum	468	23	82.1	941	9	ABM91757	Abm91757 M. xanthu
396	23	82.1	638	7	ABO69211	Abo69211 Pseudomon	469	23	82.1	948	3	AAG47641	Aag47641 Arabidops
397	23	82.1	640	8	ADR86494	Adr86494 Aspergill	470	23	82.1	975	5	ABG96290	Abg96290 Human ova
398	23	82.1	641	6	ABM68514	Abm68514 Photorhab	471	23	82.1	992	5	ABP29542	Abp29542 Streptoco
399	23	82.1	641	6	ABJ26298	Abj26298 Aspergill	472	23	82.1	992	6	ABU46823	Abu46823 Protein e
400	23	82.1	645	2	AAW22780	Aaw22780 human RAD	473	23	82.1	1002	8	ADR66344	Adr66344 Human pro
401	23	82.1	645	2	AAW71293	Aaw71293 Protein G	474	23	82.1	1002	8	ADR66686	Adr66686 Human pro
402	23	82.1	652	4	ABB60647	Abb60647 Drosophil	475	23	82.1	1008	8	ADN22976	Adn22976 Bacterial
403	23	82.1	659	7	ADF04463	Adf04463 Bacterial	476	23	82.1	1008	8	ADN22977	Adn22977 Bacterial
404	23	82.1	670	6	ABM67836	Abm67836 Photorhab	477	23	82.1	1032	5	ABP27824	Abp27824 Streptoco
405	23	82.1	674	8	ADN22628	Adn22628 Bacterial	478	23	82.1	1053	8	ABO84970	Abo84970 Murine ca
406	23	82.1	675	6	ABJ25698	Abj25698 Aspergill	479	23	82.1	1071	8	ADS43853	Ads43853 Bacterial
407	23	82.1	677	7	ABO69356	Abo69356 Pseudomon	480	23	82.1	1091	8	ADU24077	Adu24077 Human cys
408	23	82.1	679	8	ADN22629	Adn22629 Bacterial	481	23	82.1	1092	4	AAM79135	Aam79135 Human pro
409	23	82.1	694	8	ADN25736	Adn25736 Bacterial	482	23	82.1	1092	4	AAM40100	Aam40100 Human pol
410	23	82.1	695	3	AAAY45097	Aay45097 Arabidops	483	23	82.1	1092	8	ABO84971	Abo84971 Human can
411	23	82.1	695	6	AAO15964	Aao15964 Arabidops	484	23	82.1	1094	4	AAM80119	Aam80119 Human pro
412	23	82.1	695	6	AAO15963	Aao15963 Arabidops	485	23	82.1	1094	4	AAM41886	Aam41886 Human pol
413	23	82.1	695	9	ADV66378	Adv66378 Arabidops	486	23	82.1	1127	3	AAG47640	Aag47640 Arabidops
414	23	82.1	697	8	ADX94784	Adx94784 Plant ful	487	23	82.1	1173	9	ADY19886	Ady19886 PRO polyp
415	23	82.1	707	4	AAM23744	Aam23744 Human EST	488	23	82.1	1186	8	ADQ67807	Adq67807 Novel hum
416	23	82.1	707	4	AAM23715	Aam23715 Human EST	489	23	82.1	1205	6	AAE36060	Aae36060 Human tra
417	23	82.1	719	6	ABU10383	Abu10383 Mouse pha	490	23	82.1	1207	6	AAE36061	Aae36061 Human tra
418	23	82.1	720	5	ABP65117	Abp65117 Hypoxia-r	491	23	82.1	1312	2	AAW22775	Aaw22775 Human RAD
419	23	82.1	720	5	ABP59019	Abp59019 Human S4	492	23	82.1	1312	2	AAW71295	Aaw71295 Human hom
420	23	82.1	720	7	ADC13523	Adc13523 Human NOV	493	23	82.1	1312	9	ADY15954	Ady15954 PRO polyp
421	23	82.1	720	9	ADZ84580	Adz84580 Engulfmen	494	23	82.1	1312	9	ADY91604	Ady91604 Human pro
422	23	82.1	721	6	ABU10380	Abu10380 Human pha	495	23	82.1	1318	5	ABB77985	Abb77985 Amino aci
423	23	82.1	739	8	ADS45167	Ads45167 Bacterial	496	23	82.1	1318	7	ADJ68860	Adj68860 Human hea
424	23	82.1	749	5	ABB04872	Abb04872 LDL recep	497	23	82.1	1318	8	ADJ66509	Adj66509 RAD50 hom
425	23	82.1	761	5	AAE13544	Aae13544 Human cer	498	23	82.1	1318	8	ADS8326	Ads8326 Human pro
426	23	82.1	776	6	ADB12050	Adb12050 Alloiococ	499	23	82.1	1362	8	ADS34522	Ads34522 POSH prot
427	23	82.1	780	6	AAE34441	Aae34441 Human lip	500	23	82.1	1394	8	ADN25124	Adn25124 Bacterial
428	23	82.1	794	7	ABO64392	Abo64392 Klebsiell	501	23	82.1	1478	6	ABU25753	Abu25753 Protein e
429	23	82.1	815	8	ABM84044	Abm84044 Human dia	502	23	82.1	1661	2	AAW22602	Aaw22602 Ty lactone
430	23	82.1	818	8	ADX73834	Adx73834 Plant ful	503	23	82.1	1864	2	AAW22602	Aaw22602 Ty lactone
431	23	82.1	820	7	ADN95393	Adn95393 Human BEC	504	23	82.1	1868	8	ADX56097	Adx56097 Streptomy
432	23	82.1	821	6	ABR92110	Abr92110 Human cer	505	23	82.1	1881	6	ABP98810	Abp98810 Human str
433	23	82.1	821	7	ADE62260	Ade62260 Human pro	506	23	82.1	2147	6	AAO31009	Aao31009 Human tra
434	23	82.1	821	7	ADD47954	Add47954 Human pro	507	23	82.1	2654	8	ADK15820	Adk15820 Human ABC
435	23	82.1	821	7	ADF76337	Adf76337 Novel hum	508	23	82.1	2660	8	ADK15816	Adk15816 Human ABC
436	23	82.1	821	8	ADL70589	Adl70589 Cervical	509	23	82.1	2760	8	ADK15810	Adk15810 Human ABC
437	23	82.1	821	8	ADO19234	Ado19234 Human PRO	510	23	82.1	3264	9	ADY72594	Ady72594 A. orient
438	23	82.1	821	8	ABM80342	Abm80342 Tumour-as	511	23	82.1	4952	8	ADK15818	Adk15818 Human ABC
439	23	82.1	821	8	ADP54308	Adp54308 Human PRO	512	23	82.1	4958	8	ADK15814	Adk15814 Human ABC
440	23	82.1	821	8	ADU06455	Adu06455 Novel bro	513	23	82.1	5058	8	ADK15793	Adk15793 Human ABC
441	23	82.1	821	9	ADY14695	Ady14695 PRO polyp	514	22	78.6	10	2	AAAY22359	Aay22359 TPO recep
442	23	82.1	824	3	AAAY95293	Aay95293 Human GEF	515	22	78.6	10	3	ABAB16983	Aab16983 TPO-mimet
443	23	82.1	829	8	ADS42481	Ads42481 Bacterial	516	22	78.6	10	5	ABB72869	Abb72869 TPO-mimet
444	23	82.1	839	8	ADN11329	Adn11329 Aphis gos	517	22	78.6	10	7	ADJ73020	Adj73020 TPO mimet
445	23	82.1	843	7	ABO83863	Abo83863 Pseudomon	518	22	78.6	10	8	ADJ52655	Adj52655 CH1 delet
446	23	82.1	847	4	ABG17278	Abg17278 Novel hum	519	22	78.6	10	8	ADJ51616	Adj51616 CH1 delet
447	23	82.1	851	8	ADR10252	Adr10252 Human pro	520	22	78.6	15	2	AAAY22387	Aay22387 TPO recep
448	23	82.1	855	4	AAM78466	Aam78466 Human pro	521	22	78.6	15	5	ABP57970	Abp57970 Human CD4
449	23	82.1	855	4	AAM38741	Aam38741 Human pol	522	22	78.6	20	9	ADZ98167	Adz98167 Human ami
450	23	82.1	855	7	ADE59602	Ade59602 Human PRO	523	22	78.6	30	3	AAB38296	Aab38296 Human sec
451	23	82.1	855	8	ABM81053	Abm81053 Tumour-as	524	22	78.6	30	3	AAB38294	Aab38294 Human sec
452	23	82.1	868	7	ABM89313	Abm89313 Rice abio	525	22	78.6	32	4	AAO04838	Aao04838 Human pol
453	23	82.1	870	4	ABG30208	Abg30208 Novel hum	526	22	78.6	41	9	ABM96631	Abm96631 M. xanthu
454	23	82.1	871	8	ADY06939	Ady06939 Plant ful	527	22	78.6	46	3	AAAY44949	Aay44949 Human thr
455	23	82.1	873	4	AAM79450	Aam79450 Human pro	528	22	78.6	46	3	AAAY44948	Aay44948 Mouse thr
456	23	82.1	873	4	AAM40527	Aam40527 Human pol	529	22	78.6	48	4	AAO5200	Aao05200 Human pol
457	23	82.1	874	4	ABG17472	Abg17472 Novel hum	530	22	78.6	50	9	ADZ98172	Adz98172 Human ami
458	23	82.1	874	8	ADY24764	Ady24764 Plant ful	531	22	78.6	53	4	AAM88941	Aam88941 Human imm
459	23	82.1	874	8	ADY24763	Ady24763 Plant ful	532	22	78.6	60	7	ABR82695	Abr82695 Lbx1 home
460	23	82.1	882	8	ADS21711	Ads21711 Bacterial	533	22	78.6	69	4	AAU51338	Aau51338 Propionib
461	23	82.1	896	6	ABR43913	Abr43913 Human sod	534	22	78.6	69	6	ABM47857	Abm47857 Propionib
462	23	82.1	896	7	ADJ68814	Adj68814 Human hea	535	22	78.6	70	5	ABP40467	Abp40467 Staphyloc

536	22	78.6	70	8	ADS06011	Ads06011 Staphyloc	609	22	78.6	223	3	AAG08497	Aag08497 Arabidops
537	22	78.6	72	3	AAG57925	Aag57925 Zea mays	610	22	78.6	224	3	AAG51820	Aag51820 Arabidops
538	22	78.6	73	7	ADG30700	Adg30700 Xanthomon	611	22	78.6	225	9	AEA10738	Aea10738 FGF-5 bin
539	22	78.6	84	4	AAO10156	Aao10156 Human pol	612	22	78.6	226	3	AAG49179	Aag49179 Arabidops
540	22	78.6	87	4	AAU40924	Aau40924 Propionib	613	22	78.6	227	3	AAG08496	Aag08496 Arabidops
541	22	78.6	87	4	AAU65713	Aau65713 Propionib	614	22	78.6	227	8	ADT56060	Adt56060 Plant pol
542	22	78.6	87	6	ABM37443	Abm37443 Propionib	615	22	78.6	228	7	ADF73127	Adf73127 Xenopus t
543	22	78.6	87	6	ABM62232	Abm62232 Propionib	616	22	78.6	228	8	ADM99656	Adm99656 African c
544	22	78.6	92	8	ADR95435	Adr95435 Novel S.	617	22	78.6	231	6	ABU29713	Abu29713 Protein e
545	22	78.6	92	9	AEA59305	Aea59305 Streptoco	618	22	78.6	232	6	ABU02611	Abu02611 S. pneumo
546	22	78.6	93	8	ADN47583	Adn47583 Thermococ	619	22	78.6	232	6	ABP81483	Abp81483 Streptoco
547	22	78.6	95	4	ABG06872	Abg06872 Novel hum	620	22	78.6	232	6	ABU44309	Abu44309 Protein e
548	22	78.6	97	5	ABP42958	Abp42958 Human ova	621	22	78.6	232	6	ABU46283	Abu46283 Protein e
549	22	78.6	99	7	ABO60953	Ab060953 Klebsiell	622	22	78.6	232	8	ADK47700	Adk47700 Streptoco
550	22	78.6	100	3	AAG57924	Aag57924 Zea mays	623	22	78.6	234	9	AEA49217	Aea49217 L. rhamno
551	22	78.6	100	5	ABU05494	Abu05494 M. tuberc	624	22	78.6	235	3	AAG33167	Aag33167 Zea mays
552	22	78.6	100	5	ABU05471	Abu05471 M. tuberc	625	22	78.6	235	5	ABB48744	Abb48744 Listeria
553	22	78.6	100	7	ADB74329	Adb74329 Mycobacte	626	22	78.6	236	6	ABU32846	Abu32846 Protein e
554	22	78.6	101	7	ADB74435	Adb74435 Mycobacte	627	22	78.6	239	6	ABU38494	Abu38494 Protein e
555	22	78.6	105	3	AAy44960	Aay44960 KDEL rece	628	22	78.6	240	4	ABB63201	Abb63201 Drosophil
556	22	78.6	109	3	AAy44961	Aay44961 KDEL rece	629	22	78.6	244	7	ADH87382	Adh87382 Enterococ
557	22	78.6	109	3	AAy44965	Aay44965 KDEL rece	630	22	78.6	244	7	ABO72026	Abo72026 Pseudomon
558	22	78.6	114	7	ABM85732	Abm85732 Mouse pro	631	22	78.6	252	6	ABU35766	Abu35766 Protein e
559	22	78.6	116	5	ABP08252	Abp08252 Human ORF	632	22	78.6	256	5	ABB53587	Abb53587 Lactococc
560	22	78.6	116	6	ABP78249	Abp78249 N. gonorr	633	22	78.6	257	4	AAU67890	Aau67890 Propionib
561	22	78.6	117	8	ADP90619	Adp90619 Micromono	634	22	78.6	257	6	ABM64409	Abm64409 Propionib
562	22	78.6	120	8	ADX94476	Adx94476 Plant ful	635	22	78.6	266	2	AAR91940	Aar91940 Human hip
563	22	78.6	120	8	ADX90986	Adx90986 Plant ful	636	22	78.6	266	2	AAW26705	Aaw26705 Human Cyp
564	22	78.6	122	5	ABP57969	Abp57969 Human CD4	637	22	78.6	267	7	ADC96056	Adc96056 E. faeciu
565	22	78.6	125	4	AAO03224	Aao03224 Human pol	638	22	78.6	267	8	ADK52135	Adk52135 Mouse ato
566	22	78.6	134	4	AAg82640	Aag82640 S. epider	639	22	78.6	270	6	ABU23259	Abu23259 Protein e
567	22	78.6	136	3	AAG33199	Aag33199 Zea mays	640	22	78.6	271	7	ADL72407	Adl72407 Maize lon
568	22	78.6	136	3	AAG26815	Aag26815 Zea mays	641	22	78.6	277	3	AAG33166	Aag33166 Zea mays
569	22	78.6	136	4	AAO00478	Aao00478 Human pol	642	22	78.6	277	9	ADZ17581	Adz17581 Theileria
570	22	78.6	137	4	ABB67359	Abb67359 Drosophil	643	22	78.6	278	5	AAE16769	Aae16769 Human tra
571	22	78.6	138	3	AAG33233	Aag33233 Zea mays	644	22	78.6	278	7	AAO24036	Aao24036 Aspergill
572	22	78.6	138	8	ADY05350	Ady05350 Plant ful	645	22	78.6	278	8	ADU63798	Adu63798 A. Oryzae
573	22	78.6	141	3	AAB40305	Aab40305 Human ORF	646	22	78.6	278	9	ADW23376	Adw23376 Aspergill
574	22	78.6	141	5	ABP02707	Abp02707 Human ORF	647	22	78.6	290	3	AAG33165	Aag33165 Zea mays
575	22	78.6	147	5	ABP38049	Abp38049 Staphyloc	648	22	78.6	294	6	ABP78112	Abp78112 N. gonorr
576	22	78.6	147	8	ADS07409	Ads07409 Staphyloc	649	22	78.6	294	6	ABU37516	Abu37516 Protein e
577	22	78.6	149	4	AAU37961	Aau37961 Streptoco	650	22	78.6	294	8	ADL05985	Adl05985 M. catarr
578	22	78.6	153	8	ADV82718	Adv82718 Streptoco	651	22	78.6	294	8	ADT57068	Adt57068 Plant pol
579	22	78.6	154	6	ABU48752	Abu48752 Protein e	652	22	78.6	294	8	ADf45518	Adf45518 Pan trogl
580	22	78.6	159	6	AAE34015	Aae34015 Lolium pe	653	22	78.6	298	8	ADP29619	Adp29619 Human sec
581	22	78.6	160	6	ABU01398	Abu01398 S. pneumo	654	22	78.6	298	8	ADS28636	Ads28636 Bacterial
582	22	78.6	160	8	ADK48691	Adk48691 Streptoco	655	22	78.6	307	7	ADC95625	Adc95625 E. faeciu
583	22	78.6	161	3	AAG15656	Aag15656 Arabidops	656	22	78.6	309	4	AAG70725	Aag70725 S. cerevis
584	22	78.6	161	4	ABB52923	Abb52923 Escherich	657	22	78.6	309	6	ADB10430	Adb10430 Alloiococ
585	22	78.6	161	4	ABB52957	Abb52957 Escherich	658	22	78.6	309	8	ADS43514	Ads43514 Bacterial
586	22	78.6	161	8	ADR94930	Adr94930 Novel S.	659	22	78.6	310	4	AAG72206	Aag72206 Human Olf
587	22	78.6	161	9	AEA58800	Aea58800 Streptoco	660	22	78.6	310	4	AAE04582	Aae04582 Human G-p
588	22	78.6	173	3	AAG33198	Aag33198 Zea mays	661	22	78.6	310	5	AAU95560	Aau95560 Human Olf
589	22	78.6	175	9	ADX02798	Adx02798 Sendai vi	662	22	78.6	310	5	AAU83574	Aau83574 Human nov
590	22	78.6	176	8	ADR75253	Adr75253 Mouse gas	663	22	78.6	310	5	AAU83573	Aau83573 Human nov
591	22	78.6	181	9	ADX02797	Adx02797 Sendai vi	664	22	78.6	310	5	AAU83574	Aau83574 Human nov
592	22	78.6	182	2	AAy35544	Aay35544 C. pneumo	665	22	78.6	310	7	ADC86081	Adc86081 Human GPC
593	22	78.6	200	4	AAU54511	Aau54511 Propionib	666	22	78.6	310	7	ADM29630	Adm29630 Novel hum
594	22	78.6	200	6	ABM51030	Abm51030 Propionib	667	22	78.6	310	7	ADM29628	Adm29628 Novel hum
595	22	78.6	203	1	AAp70394	Aap70394 Sequence	668	22	78.6	314	4	ABB59817	Abb59817 Drosophil
596	22	78.6	204	7	ADC85520	Adc85520 Sendai vi	669	22	78.6	315	3	AAG32359	Aag32359 Arabidops
597	22	78.6	204	9	ADX02796	Adx02796 Sendai vi	670	22	78.6	316	5	ADI16828	Adi16828 Murine NO
598	22	78.6	205	8	ADS24392	Ads24392 Bacterial	671	22	78.6	316	6	ABO19500	Abo19500 Mouse GPC
599	22	78.6	207	4	ABG13866	Abg13866 Novel hum	672	22	78.6	316	7	ADF71450	Adf71450 Wheat pan
600	22	78.6	214	4	AAG72857	Aag72857 Human Olf	673	22	78.6	316	7	ABW02145	Abw02145 Human GPC
601	22	78.6	215	9	ADX02795	Adx02795 Sendai vi	674	22	78.6	316	8	ADQ98158	Adq98158 Mouse Olf
602	22	78.6	218	4	AAG90916	Aag90916 C glutami	675	22	78.6	316	9	ADZ39245	Adz39245 Wheat pan
603	22	78.6	218	9	AEB15287	Aeb15287 C glutami	676	22	78.6	317	3	AAG17047	Aag17047 Arabidops
604	22	78.6	221	3	AAG32360	Aag32360 Arabidops	677	22	78.6	318	3	AAG51819	Aag51819 Arabidops
605	22	78.6	222	5	ABB48132	Abb48132 Listeria	678	22	78.6	319	8	ADY08458	Ady08458 Plant ful
606	22	78.6	222	6	ABU32818	Abu32818 Protein e	679	22	78.6	320	7	ADF06693	Adf06693 Bacterial
607	22	78.6	223	3	AAG17048	Aag17048 Arabidops	680	22	78.6	322	4	AAU47175	Aau47175 Propionib
608	22	78.6	223	3	AAG49180	Aag49180 Arabidops	681	22	78.6	322	6	ABM70262	Abm70262 Photorhab

682	22	78.6	322	6	ABM43694	Abm43694 Propionib
683	22	78.6	324	4	ABB70055	Abb70055 Drosophil
684	22	78.6	326	4	AAG72856	Aag72856 Human olf
685	22	78.6	328	3	AAG06187	Aag06187 Arabidops
686	22	78.6	328	3	AAG49302	Aag49302 Arabidops
687	22	78.6	328	8	ADX70908	Adx70908 Plant ful
688	22	78.6	330	6	ABU21137	Abu21137 Protein e
689	22	78.6	331	5	ABP73685	Abp73685 Candida a
690	22	78.6	332	4	ABG28228	Abg28228 Novel hum
691	22	78.6	332	6	ABU49437	Abu49437 Protein e
692	22	78.6	333	6	ABU47009	Abu47009 Protein e
693	22	78.6	333	8	ADQ37063	Adq37063 Cell prol
694	22	78.6	337	3	AAG51653	Aag51653 Arabidops
695	22	78.6	340	9	AEA79823	Aea79823 BCE103 ce
696	22	78.6	342	6	ABM69247	Abm69247 Photorhab
697	22	78.6	343	3	AAB32510	Aab32510 S. lavend
698	22	78.6	343	7	ADE10282	Adel0282 S. lavend
699	22	78.6	344	3	AAG10368	Aag10368 Arabidops
700	22	78.6	344	3	AAG45683	Aag45683 Arabidops
701	22	78.6	345	6	ABU17023	Abu17023 Protein e
702	22	78.6	346	4	ABG02100	Abg02100 Novel hum
703	22	78.6	347	4	AAG72620	Aag72620 Murine OR
704	22	78.6	349	6	ABU44697	Abu44697 Protein e
705	22	78.6	350	6	ADA33389	Ada33389 Acinetoba
706	22	78.6	350	8	ADS43305	Ads43305 Bacterial
707	22	78.6	355	4	ABG28230	Abg28230 Novel hum
708	22	78.6	356	6	ABM68580	Abm68580 Photorhab
709	22	78.6	356	6	ABM69316	Abm69316 Photorhab
710	22	78.6	356	6	ABM68997	Abm68997 Photorhab
711	22	78.6	356	6	ABM67343	Abm67343 Photorhab
712	22	78.6	356	6	ABM67680	Abm67680 Photorhab
713	22	78.6	356	6	ABM68134	Abm68134 Photorhab
714	22	78.6	356	6	ABM68359	Abm68359 Photorhab
715	22	78.6	358	7	ABM86427	Abm86427 Rice abio
716	22	78.6	360	4	AAU24527	Aau24527 Human olf
717	22	78.6	360	5	AAU85151	Aau85151 G-coupled
718	22	78.6	366	8	ADN23745	Adn23745 Bacterial
719	22	78.6	369	4	AAU50912	Aau50912 Propionib
720	22	78.6	369	6	ABM47431	Abm47431 Propionib
721	22	78.6	370	4	AAB79723	Aab79723 Corynebac
722	22	78.6	370	4	AAB79728	Aab79728 Corynebac
723	22	78.6	370	4	AAU71899	Aau71899 C. glutam
724	22	78.6	372	6	ADA34058	Ada34058 Acinetoba
725	22	78.6	374	8	ADU47402	Adu47402 Corynebac
726	22	78.6	377	4	AAB79722	Aab79722 Corynebac
727	22	78.6	377	4	AAB79727	Aab79727 Corynebac
728	22	78.6	377	4	AAU71898	Aau71898 C. glutam
729	22	78.6	377	4	AAG90471	Aag90471 C. glutami
730	22	78.6	378	8	ADJ49740	Adj49740 Oil-assoc
731	22	78.6	379	8	ADO80279	Ado80279 Mycobacte
732	22	78.6	381	4	AAB52461	Aab52461 Mycobacte
733	22	78.6	381	7	ADC60855	Adc60855 M. tuberc
734	22	78.6	382	6	ADB07806	Adb07806 Alloiococ
735	22	78.6	382	8	ADO80277	Ado80277 Mycobacte
736	22	78.6	385	2	AAJ39314	Aay39314 SpnR prot
737	22	78.6	385	3	AAG49301	Aag49301 Arabidops
738	22	78.6	385	3	AAG06186	Aag06186 Arabidops
739	22	78.6	385	4	AAB70961	Aab70961 S. spinos
740	22	78.6	385	9	ADZ99205	Adz99205 S. spinos
741	22	78.6	385	9	ADZ99291	Adz99291 S. spinos
742	22	78.6	385	9	ADZ99299	Adz99299 S. spinos
743	22	78.6	385	9	AEA10735	Aea10735 FGF-5 bin
744	22	78.6	385	9	AEA10742	Aea10742 FGF-5 bin
745	22	78.6	385	9	AEA10731	Aea10731 FGF-5 bin
746	22	78.6	385	9	AEA10733	Aea10733 FGF-5 bin
747	22	78.6	385	9	AEA10740	Aea10740 FGF-5 bin
748	22	78.6	385	9	AEA10727	Aea10727 FGF-5 bin
749	22	78.6	385	9	AEA10729	Aea10729 FGF-5 bin
750	22	78.6	386	3	AAG32358	Aag32358 Arabidops
751	22	78.6	387	6	ABU45056	Abu45056 Protein e
752	22	78.6	387	7	ADJ70100	Adj70100 Human hea
753	22	78.6	387	8	ADO62666	Ado62666 Transcrip
754	22	78.6	388	3	AAG17046	Aag17046 Arabidops

755	22	78.6	388	5	ABB93126	Abb93126 Herbicida
756	22	78.6	388	8	ADS24489	Ads24489 Bacterial
757	22	78.6	389	3	AAG51818	Aag51818 Arabidops
758	22	78.6	391	3	AAG06185	Aag06185 Arabidops
759	22	78.6	391	3	AAG49300	Aag49300 Arabidops
760	22	78.6	391	5	ABB93709	Abb93709 Herbicida
761	22	78.6	391	6	ABU47379	Abu47379 Protein e
762	22	78.6	392	8	ADY05588	Ady05588 Plant ful
763	22	78.6	395	4	AAU38299	Aau38299 Salmonell
764	22	78.6	398	8	ADS22029	Ads22029 Bacterial
765	22	78.6	398	8	ADX74065	Adx74065 Plant ful
766	22	78.6	398	8	ADX92502	Adx92502 Plant ful
767	22	78.6	399	2	AAR74660	Aar74660 Pseudomon
768	22	78.6	400	2	AAW22521	Aaw22521 Bacillus
769	22	78.6	400	2	AAW23601	Aaw23601 Bacillus
770	22	78.6	400	2	AAW57431	Aaw57431 Bacillus
771	22	78.6	401	6	ABU19990	Abu19990 Protein e
772	22	78.6	403	6	ABU49552	Abu49552 Protein e
773	22	78.6	404	6	ADB07808	Adb07808 Alloiococ
774	22	78.6	406	8	ADY04935	Ady04935 Plant ful
775	22	78.6	407	6	ABU40512	Abu40512 Protein e
776	22	78.6	407	8	ADX93230	Adx93230 Plant ful
777	22	78.6	409	1	AAP81843	Aap81843 Sequence
778	22	78.6	410	2	AAW12378	Aaw12378 P300-CelB
779	22	78.6	410	6	ADB12347	Adb12347 Alloiococ
780	22	78.6	411	2	AAW12379	Aaw12379 P300-CelB
781	22	78.6	411	2	AAW12381	Aaw12381 P300-CelB
782	22	78.6	412	2	AAW12380	Aaw12380 P300-CelB
783	22	78.6	412	6	ABU33108	Abu33108 Protein e
784	22	78.6	412	9	AEBA1502	Aeb41502 L. pneumo
785	22	78.6	413	8	ADX77667	Adx77667 Plant ful
786	22	78.6	415	7	ABM89908	Abm89908 Rice abio
787	22	78.6	416	6	ABU30633	Abu30633 Protein e
788	22	78.6	416	7	ADF06308	Adf06308 Bacterial
789	22	78.6	417	8	ADN25102	Adn25102 Bacterial
790	22	78.6	420	6	ABU39017	Abu39017 Protein e
791	22	78.6	424	3	AAG29585	Aag29585 Arabidops
792	22	78.6	424	8	ADJ56879	Adj56879 Arabidops
793	22	78.6	424	9	AEBA38215	Aeb38215 L. pneumo
794	22	78.6	425	3	AAG29584	Aag29584 Arabidops
795	22	78.6	425	8	ADX79792	Adx79792 Plant ful
796	22	78.6	426	3	AAG13338	Aag13338 Arabidops
797	22	78.6	426	3	AAG51652	Aag51652 Arabidops
798	22	78.6	428	9	AEA29532	Aea29532 Fusion pr
799	22	78.6	429	6	ABU47820	Abu47820 Protein e
800	22	78.6	429	6	ABU47200	Abu47200 Protein e
801	22	78.6	430	3	AAG29583	Aag29583 Arabidops
802	22	78.6	431	6	ADB12085	Adb12085 Alloiococ
803	22	78.6	432	6	ADB08864	Adb08864 Alloiococ
804	22	78.6	433	6	ADB12345	Adb12345 Alloiococ
805	22	78.6	433	8	ADJ49755	Adj49755 Oil-assoc
806	22	78.6	433	8	ADS22912	Ads22912 Bacterial
807	22	78.6	435	3	AAJ75574	Aay75574 Neisseria
808	22	78.6	440	6	ABM70591	Abm70591 Photorhab
809	22	78.6	443	8	ADX91485	Adx91485 Plant ful
810	22	78.6	444	8	ADS24875	Ads24875 Bacterial
811	22	78.6	447	4	AAG74174	Aag74174 Human col
812	22	78.6	447	6	ABU22101	Abu22101 Protein e
813	22	78.6	447	8	ADJ49780	Adj49780 Oil-assoc
814	22	78.6	448	8	ADY11816	Ady11816 Plant ful
815	22	78.6	451	5	ABP69683	Abp69683 Human pol
816	22	78.6	455	4	ABB11613	Abb11613 Human sor
817	22	78.6	457	8	ADN25370	Adn25370 Bacterial
818	22	78.6	460	7	ABO79022	Abo79022 Pseudomon
819	22	78.6	461	5	ABB05724	Abb05724 Human tes
820	22	78.6	462	2	AAW57433	Aaw57433 Cloned al
821	22	78.6	467	2	AAW05731	Aaw05731 Cellulase
822	22	78.6	467	2	AAW00382	Aaw00382 Bacillus
823	22	78.6	467	5	ABB81990	Abb81990 Bacillus
824	22	78.6	467	6	ABU17917	Abu17917 Protein e
825	22	78.6	469	8	ADN60465	Adn60465 B. lichen
826	22	78.6	470	6	ABU17011	Abu17011 Protein e
827	22	78.6	471	9	AEBA1489	Aeb41489 L. pneumo

828	22	78.6	472	6	ABU18104	Abul18104 Protein e	901	22	78.6	584	8	ADS28133	Ads28133 Bacterial
829	22	78.6	473	5	ABG32151	Abg32151 Arabidops	902	22	78.6	592	7	ABO64264	AbO64264 Klebsiell
830	22	78.6	473	9	ADY64606	Aay64606 S. manson	903	22	78.6	592	9	AEA43084	Aea43084 M. smegma
831	22	78.6	473	3	AAy75573	Aay75573 Neisseria	904	22	78.6	592	9	AEA43085	Aea43085 M. smegma
832	22	78.6	475	6	AAE36320	Aae36320 Moraxella	905	22	78.6	593	8	ADS14851	Adsi14851 Pseudomon
833	22	78.6	477	6	ADB10428	Adb10428 Alloiococ	906	22	78.6	595	7	ABO78285	AbO78285 Pseudomon
834	22	78.6	481	6	ADA33285	Ada33285 Acinetoba	907	22	78.6	605	7	ADJ49335	Adj49335 Oil-assoc
835	22	78.6	482	4	AAU38270	Aau38270 Salmonell	908	22	78.6	616	7	ABO79391	AbO79391 Pseudomon
836	22	78.6	484	8	ADN18452	Adn18452 Bacterial	909	22	78.6	618	5	ABP26565	Abp26565 Streptoco
837	22	78.6	485	8	ADU07677	Adu07677 Amino aci	910	22	78.6	618	6	ABU46844	Abu46844 Protein e
838	22	78.6	485	9	ABE38200	Aeb38200 L. pneumo	911	22	78.6	619	7	ADH87640	Adh87640 Enterococ
839	22	78.6	488	3	AAG51651	Aag51651 Arabidops	912	22	78.6	624	8	ADX97030	Adx97030 Plant ful
840	22	78.6	488	3	AAG13337	Aag13337 Arabidops	913	22	78.6	629	7	ABO77217	AbO77217 Pseudomon
841	22	78.6	488	4	ABG60032	Abg60032 Drosophil	914	22	78.6	636	5	AAM49536	Aam49536 S. cynthi
842	22	78.6	488	5	ABG32153	Abg32153 Arabidops	915	22	78.6	648	7	ABO71382	AbO71382 Pseudomon
843	22	78.6	488	7	ADB95062	Adb95062 A. thalia	916	22	78.6	652	5	ABP66325	Abp66325 Bifidobac
844	22	78.6	488	8	ADH38674	Adh38674 Aarbisops	917	22	78.6	660	8	ADN22630	Adn22630 Bacterial
845	22	78.6	488	9	ADZ00121	Adz00121 Arabidops	918	22	78.6	661	2	AAR26176	Aar26176 GORgab pr
846	22	78.6	489	8	ADX87377	Adx87377 Plant ful	919	22	78.6	661	2	AAR22674	Aar22674 Recombina
847	22	78.6	495	4	ABG17680	Abg17680 Novel hum	920	22	78.6	662	8	ADN22631	Adn22631 Bacterial
848	22	78.6	497	5	AAE22310	Aae22310 Methylomo	921	22	78.6	673	4	ABB59262	Abb59262 Drosophil
849	22	78.6	497	5	ABG61589	Abg61589 High grow	922	22	78.6	673	7	ABO79541	AbO79541 Pseudomon
850	22	78.6	497	5	AAU80333	Aau80333 Methylomo	923	22	78.6	675	4	ABG16318	Abg16318 Novel hum
851	22	78.6	497	6	ADA14538	Ada14538 Methylomo	924	22	78.6	675	8	ADS12244	Adsl2244 Human the
852	22	78.6	497	9	ADW08985	Adw08985 Methylomo	925	22	78.6	679	7	ABO80141	AbO80141 Pseudomon
853	22	78.6	497	9	ABM94498	Abm94498 M. xanthu	926	22	78.6	690	4	ABB66108	Abb66108 Drosophil
854	22	78.6	504	8	ADJ49680	Adj49680 Oil-assoc	927	22	78.6	690	4	ABB61438	Abb61438 Drosophil
855	22	78.6	504	8	ADZ00132	Adz00132 A. thalia	928	22	78.6	696	7	ABO71841	AbO71841 Pseudomon
856	22	78.6	511	8	ADY07023	Ady07023 Plant ful	929	22	78.6	699	8	ADI26806	Adi26806 Saccharom
857	22	78.6	514	5	AAU80014	Aau80014 Confiral	930	22	78.6	709	6	ABU40849	Abu40849 Protein e
858	22	78.6	514	7	ADD93898	Add93898 Quaking a	931	22	78.6	712	6	ABJ25450	Abj25450 Aspergill
859	22	78.6	517	4	AAU60053	Aau60053 Propionib	932	22	78.6	716	8	ADL06147	Adl06147 M. catarr
860	22	78.6	517	6	ABM56582	Abm56582 Propionib	933	22	78.6	718	4	ABB66573	Abb66573 Drosophil
861	22	78.6	520	8	ADS42561	Ads42561 Bacterial	934	22	78.6	719	7	ABM86452	Abm86452 Rice abio
862	22	78.6	521	8	ADS27891	Ads27891 Bacterial	935	22	78.6	719	7	ABM90154	Abm90154 Rice abio
863	22	78.6	522	8	ADN20865	Adn20865 Bacterial	936	22	78.6	720	7	ADF05453	Adf05453 Bacterial
864	22	78.6	525	9	ADZ46904	Adz46904 BASB232 p	937	22	78.6	724	4	ABB61429	Abb61429 Drosophil
865	22	78.6	526	8	ADS29015	Ads29015 Bacterial	938	22	78.6	724	6	ABU10378	Abu10378 Drosophil
866	22	78.6	527	3	AAG13336	Aag13336 Arabidops	939	22	78.6	735	8	ADS43821	Ads43821 Bacterial
867	22	78.6	529	3	AAG43981	Aag43981 Zea mays	940	22	78.6	736	6	ABU61998	Abu61998 Human thr
868	22	78.6	530	2	AAR12276	Aar12276 Recombina	941	22	78.6	736	8	ADR86001	Adr86001 Aspergill
869	22	78.6	530	6	ABU31766	Abu31766 Protein e	942	22	78.6	755	7	ADK63620	Adk63620 Disease t
870	22	78.6	533	4	AAU37798	Aau37798 Streptoco	943	22	78.6	755	8	ADN19332	Adn19332 Bacterial
871	22	78.6	533	4	AAm01018	Aam01018 CFE 17 pr	944	22	78.6	762	4	AAM93335	Aam93335 Human pol
872	22	78.6	533	5	ABP26773	Abp26773 Streptoco	945	22	78.6	762	8	ADL30836	Adl30836 Human pro
873	22	78.6	533	6	ABU00989	Abu00989 S. pneumo	946	22	78.6	774	8	ADR86079	Adr86079 Aspergill
874	22	78.6	533	6	ABU45930	Abu45930 Protein e	947	22	78.6	782	7	ADD48693	Add48693 Rat Prote
875	22	78.6	533	8	ADK48791	Adk48791 Streptoco	948	22	78.6	782	8	ADI82106	Adi82106 Fibrinoge
876	22	78.6	533	8	ADT50098	Adt50098 S.pneumon	949	22	78.6	796	6	ABR53413	Abr53413 Protein s
877	22	78.6	533	8	ADV89309	Adv89309 Streptoco	950	22	78.6	796	7	ADK64704	Adk64704 Disease t
878	22	78.6	533	8	ADV82733	Adv82733 Streptoco	951	22	78.6	817	7	ADE08397	Ade08397 Novel pro
879	22	78.6	533	8	ADV80562	Adv80562 Streptoco	952	22	78.6	837	5	ADF45519	Adf45519 Rat KIAA1
880	22	78.6	554	6	ABU44198	Abu44198 Protein e	953	22	78.6	841	8	ABP68967	Abp68967 Human pol
881	22	78.6	555	5	ABB47951	Abb47951 Listeria	954	22	78.6	841	8	ADF45517	Adf45517 Human KIA
882	22	78.6	555	6	ABU32591	Abu32591 Protein e	955	22	78.6	841	8	ADF45492	Adf45492 Human KIA
883	22	78.6	557	8	ADS15044	Ads15044 Pseudomon	956	22	78.6	845	7	ABO69110	AbO69110 Pseudomon
884	22	78.6	558	8	ADR95837	Adr95837 Novel S.	957	22	78.6	872	5	ABU27040	Abu27040 Protein e
885	22	78.6	558	8	ADX95129	Adx95129 Plant ful	958	22	78.6	875	5	ABU05813	Abu05813 M. tuberc
886	22	78.6	558	9	AEA59707	Aea59707 Streptoco	959	22	78.6	875	6	ABU36595	Abu36595 Protein e
887	22	78.6	561	4	AAU36080	Aau36080 Klebsiell	960	22	78.6	875	6	ABU34905	Abu34905 Protein e
888	22	78.6	562	6	ABU29789	Abu29789 Protein e	961	22	78.6	880	8	ADY22311	Ady22311 Plant ful
889	22	78.6	564	7	ADC95621	Adc95621 E. faeciu	962	22	78.6	884	6	ABU34793	Abu34793 Protein e
890	22	78.6	565	3	AAG43980	Aag43980 Zea mays	963	22	78.6	884	6	ABU36818	Abu36818 Protein e
891	22	78.6	575	6	ABU47060	Abu47060 Protein e	964	22	78.6	891	8	ADX93197	Adx93197 Plant ful
892	22	78.6	575	6	ABU31409	Abu31409 Protein e	965	22	78.6	901	9	AEB17427	Aeb17427 E. coli K
893	22	78.6	577	2	AAW85704	Aaw85704 Grand Fir	966	22	78.6	903	7	ADN20859	Adn20859 Bacterial
894	22	78.6	577	4	AAE69374	Aae69374 Grand fir	967	22	78.6	927	7	ABO64572	AbO64572 Klebsiell
895	22	78.6	578	7	ABO62639	AbO62639 Klebsiell	968	22	78.6	928	8	ADS42628	Ads42628 Bacterial
896	22	78.6	580	2	AAy06571	Aay06571 Delta-sel	969	22	78.6	931	6	ABU33399	Abu33399 Protein e
897	22	78.6	581	2	AAy06570	Aay06570 Delta-sel	970	22	78.6	931	9	AEB39466	Aeb39466 L. pneumo
898	22	78.6	581	2	AAy06563	Aay06563 Grand fir	971	22	78.6	932	7	ABM90297	Abm90297 Rice abio
899	22	78.6	581	2	AAy06569	Aay06569 Delta-sel	972	22	78.6	933	9	AEB36038	Aeb36038 L. pneumo
900	22	78.6	581	3	AAy90854	Aay90854 Grand fir	973	22	78.6	941	8	ABM84487	Abm84487 Human dia

974 22 78.6 949 8 ABM84486 Abm84486 Human dia
975 22 78.6 956 7 ADP65204 Adp65204 Human thr
976 22 78.6 956 8 ADQ19408 Adq19408 Human sof
977 22 78.6 957 8 ADS27052 Ads27052 Bacterial
978 22 78.6 958 8 ABM84485 Abm84485 Human dia
979 22 78.6 960 2 AAY29797 Aay29797 Murine ga
980 22 78.6 960 4 AAB50089 Aab50089 Murine GA
981 22 78.6 960 5 ABB05633 Abb05633 Murine GA
982 22 78.6 960 6 ABP70633 Abp70633 Amino aci
983 22 78.6 960 6 ABM04786 Abm04786 Murine GA
984 22 78.6 963 8 ADS26661 Ads26661 Bacterial
985 22 78.6 968 8 ABM84484 Abm84484 Human dia
986 22 78.6 1031 9 ABM93606 Abm93606 M. xanthu
987 22 78.6 1044 8 ADT55557 Adt55557 Plant pol
988 22 78.6 1052 7 ABM89794 Abm89794 Rice abio
989 22 78.6 1056 3 AAG31894 Aag31894 Arabidops
990 22 78.6 1073 3 AAG31893 Aag31893 Arabidops
991 22 78.6 1079 3 AAG48607 Aag48607 Arabidops
992 22 78.6 1087 3 AAG31892 Aag31892 Arabidops
993 22 78.6 1087 7 ADB67081 Adb67081 Human A-k
994 22 78.6 1087 8 ADF45529 Adf45529 Rabbit AK
995 22 78.6 1087 8 ADL99352 Adl99352 Nanostruc
996 22 78.6 1096 3 AAG48606 Aag48606 Arabidops
997 22 78.6 1104 8 ADN21482 Adn21482 Bacterial
998 22 78.6 1110 3 AAG48605 Aag48605 Arabidops
999 22 78.6 1120 7 ADJ72235 Adj72235 S roseosp
1000 22 78.6 1238 4 ABB62022 Abb62022 Drosophil

ALIGNMENTS

RESULT 1
AAY22365
ID AAY22365 standard; peptide; 11 AA.
XX
AC AAY22365;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 16.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 11 AA;
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 2; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
Db 3 VRDQIMLSL 11
RESULT 2
AAY22370
ID AAY22370 standard; peptide; 11 AA.
XX
AC AAY22370;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 21.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 2; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9

Db 3 VRDQIWAAL 11

RESULT 3
AAB16986
ID AAB16986 standard; peptide; 11 AA.
XX
AC AAB16986;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:42.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 210; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 3 VRDQIWAAL 11

RESULT 5
ABB72872

RESULT 4
AAB16988
ID AAB16988 standard; peptide; 11 AA.
XX
AC AAB16988;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:44.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 210; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 3 VRDQIMLSL 11

RESULT 5
ABB72872

ID ABB72872 standard; peptide; 11 AA.
XX ABB72872;
AC
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:42.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytosstatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytosstatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;

Query Match 92.9%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||| |
Db 3 VRDQIWAAL 11

RESULT 6
ABB72874
ID ABB72874 standard; peptide; 11 AA.
XX
AC ABB72874;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:44.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytosstatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytosstatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention


```
XX
SQ      Sequence 11 AA;

Query Match      92.9%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VRDQIXXXL 9
Db      3 VRDQIMLSL 11

RESULT 7
ADJ73023
ID      ADJ73023 standard; peptide; 11 AA.
XX
AC      ADJ73023;
XX
DT      06-MAY-2004 (first entry)
XX
DE      TPO mimetic peptide sequence SeqID 477.
XX
KW      mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW      cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW      immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW      TPO.
XX
OS      Synthetic.
XX
PN      WO2003084477-A2.
XX
PD      16-OCT-2003.
XX
PF      24-MAR-2003; 2003WO-US009139.
XX
PR      29-MAR-2002; 2002US-0368791P.
XX
PA      (CENZ ) CENTOCOR INC.
XX
PI      Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR      WPI; 2003-804237/75.
XX
PT      New CDR mimetibody comprising a portion of a heavy or light chain
PT      variable region comprising human framework or ligand binding region,
PT      useful for preparing a composition for treating e.g., immune,
PT      cardiovascular or neurologic disease.
XX
PS      Disclosure; SEQ ID NO 477; 97pp; English.
XX
CC      This invention relates to novel mammalian CDR mimetibodies, specific
CC      portions or variants thereof. Specifically, it refers to an antibody
CC      fragment where a protein has been inserted into, or replaces a portion
CC      of, one or more CDR regions, such that each CDR mimetibody comprises at
CC      least one portion of a heavy chain or light chain variable region, which
CC      itself comprises at least one human framework region and at least one
CC      ligand binding region (LBR). The present invention describes human
CC      mimetibodies, including modified immunoglobulins and cleavage products
CC      that can be useful in gene therapy and the generation of transgenic
CC      plants and animals. Furthermore, the CDR mimetibody is useful for
CC      preparing compositions for modulating, treating or reducing the symptoms
CC      of immune, cardiovascular, infectious, malignant and/ or neurologic
CC      diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC      cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC      peptide sequence is a TPO mimetic peptide sequence used to make a
CC      mimetibody of the invention.
XX
SQ      Sequence 11 AA;

Query Match      92.9%; Score 26; DB 7; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX
SQ      Sequence 11 AA;

Query Match      92.9%; Score 26; DB 7; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 VRDQIXXXL 9
Db      3 VRDQIWAAL 11

RESULT 8
ADJ73025
ID      ADJ73025 standard; peptide; 11 AA.
XX
AC      ADJ73025;
XX
DT      06-MAY-2004 (first entry)
XX
DE      TPO mimetic peptide sequence SeqID 479.
XX
KW      mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW      cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW      immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW      TPO.
XX
OS      Synthetic.
XX
PN      WO2003084477-A2.
XX
PD      16-OCT-2003.
XX
PF      24-MAR-2003; 2003WO-US009139.
XX
PR      29-MAR-2002; 2002US-0368791P.
XX
PA      (CENZ ) CENTOCOR INC.
XX
PI      Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR      WPI; 2003-804237/75.
XX
PT      New CDR mimetibody comprising a portion of a heavy or light chain
PT      variable region comprising human framework or ligand binding region,
PT      useful for preparing a composition for treating e.g., immune,
PT      cardiovascular or neurologic disease.
XX
PS      Disclosure; SEQ ID NO 479; 97pp; English.
XX
CC      This invention relates to novel mammalian CDR mimetibodies, specific
CC      portions or variants thereof. Specifically, it refers to an antibody
CC      fragment where a protein has been inserted into, or replaces a portion
CC      of, one or more CDR regions, such that each CDR mimetibody comprises at
CC      least one portion of a heavy chain or light chain variable region, which
CC      itself comprises at least one human framework region and at least one
CC      ligand binding region (LBR). The present invention describes human
CC      mimetibodies, including modified immunoglobulins and cleavage products
CC      that can be useful in gene therapy and the generation of transgenic
CC      plants and animals. Furthermore, the CDR mimetibody is useful for
CC      preparing compositions for modulating, treating or reducing the symptoms
CC      of immune, cardiovascular, infectious, malignant and/ or neurologic
CC      diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC      cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC      peptide sequence is a TPO mimetic peptide sequence used to make a
CC      mimetibody of the invention.
XX
SQ      Sequence 11 AA;

Query Match      92.9%; Score 26; DB 7; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 VRDQIXXXL 9
Db      3 VRDQIMLSL 11

RESULT 9
ADJ52658
```

ID ADJ52658 standard; peptide; 11 AA.
XX
AC ADJ52658;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID477.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 477; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 11 AA;
Query Match 92.9%; Score 26; DB 8; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
Db 3 VRDQIWAAL 11
RESULT 10
ADJ52660
ID ADJ52660 standard; peptide; 11 AA.
XX
AC ADJ52660;

XX 06-MAY-2004 (first entry)
DT
XX
DE CH1 deleted mimetibody-related peptide SeqID479.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 479; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 11 AA;
Query Match 92.9%; Score 26; DB 8; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
Db 3 VRDQIMLSL 11
RESULT 11
ADJ51619
ID ADJ51619 standard; peptide; 11 AA.
XX
AC ADJ51619;
XX
DT 06-MAY-2004 (first entry)
XX

DE CH1 deleted mimetibody-related peptide SeqID477.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunological disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

OS WO2004002424-A2.

XX 08-JAN-2004.

PN 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

XX Kutoloski KA;

PI WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for

DR diagnosing, preventing or treating cardiovascular, dermatologic,

XX endocrine, gastrointestinal, gynecologic, infectious, neurologic and

XX nutritional disorders.

PS Claim 14; SEQ ID NO 477; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CH1 deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, gynaecological, hepatic,

CC endocrine, metabolic, gastrointestinal, allergic, infectious,

CC obstetric, haematologic, immunological, nutritional, ophthalmologic,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX Sequence 11 AA;

Query Match 92.9%; Score 26; DB 8; Length 11;

Best Local Similarity 66.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

Db 3 VRDQIWAAL 11

RESULT 12

ADJ51621

ID ADJ51621 standard; peptide; 11 AA.

XX AC ADJ51621;

XX 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID479.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunological disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

OS WO2004002424-A2.

XX 08-JAN-2004.

PN 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for

PT diagnosing, preventing or treating cardiovascular, dermatologic,

PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 14; SEQ ID NO 479; 123pp; English.

PS This invention relates to CH1 deleted mimetibodies (and the DNA sequences

XX which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CH1 deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, gynaecological, hepatic,

CC endocrine, metabolic, gastrointestinal, allergic, infectious,

CC obstetric, haematologic, immunological, nutritional, ophthalmologic,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX Sequence 11 AA;

Query Match 92.9%; Score 26; DB 8; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 3 VRDQIMLSL 11

RESULT 13
AAAY22407
ID AAY22407 standard; peptide; 13 AA.
XX AC
XX AAY22407;
DT 27-SEP-1999 (first entry)
XX TPO receptor binding peptide sequence, SEQ ID NO. 58.
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX OS Synthetic.
XX US5932546-A.
PN 03-AUG-1999.
PD 04-OCT-1996; 96US-00726464.
XX 04-OCT-1996; 96US-00726464.
PA (GLAX) GLAXO WELLCOME INC.
XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirila SE, Johnson SS;
XX WPI; 1999-457122/38.
DR New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX Disclosure; Col 15; 36pp; English.
PS This sequence represents a thrombopoietin (TPO) receptor (TR) binding
XX peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration

Sequence 13 AA;

Query Match 92.9%; Score 26; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 3 VRDQIMLSL 11

RESULT 14
AAAY65258
ID AAY65258 standard; protein; 95 AA.
XX

AC AAY65258;
XX 01-FEB-2000 (first entry)
DT Human 5' EST related polypeptide SEQ ID NO:1419.
DE Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
XX gene therapy; chromosome mapping; upstream regulatory sequence; forensic;
KW location; development; protein synthesis; stability; regulation;
KW identification.
XX Homo sapiens.
OS WO9953051-A2.
XX 21-OCT-1999.
PD 09-APR-1999; 99WO-IB0000712.
PF 09-APR-1998; 98US-00057719.
XX 28-APR-1998; 98US-00069047.
PR (GEST) GENSET.
KW Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-038446/03.
PI N-PSDB; AAZ42872.
DR Novel secreted protein 5' expressed sequence tag sequences used in
XX diagnostic, forensic, gene therapy, and chromosome mapping procedures.
PT Claim 3; Page 779; 837pp; English.
PS AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
XX sequences, corresponding to human secreted proteins. AAY64651 to AAY65438
CC represent the EST-related proteins corresponding to AAZ42265 to AAZ43052.
CC The 5' ESTs can be used for producing secreted human gene products. They
CC can be used to identify and isolate 5' untranslated regions (UTRs) and
CC upstream regulatory regions which control the location, development
CC stage, rate, and quantity of protein synthesis, as well as stability of
CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to
CC obtain full length cDNA clones. The ESTs can also be used in forensic
CC procedures to identify individuals, or in diagnostic procedures to
CC identify individuals having genetic diseases resulting from abnormal gene
CC expression. The products may also be used in gene therapy protocols. The
CC nucleic acids encoding signal peptides can be used for directing
CC extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell. The
CC proteins encoded by the EST sequences may be useful in treating a variety
CC of human conditions. Secreted proteins have therapeutic value, and the
CC identification of new secreted proteins is valuable. AAZ42249 to AAZ42264
CC and AAY64644 to AAY64650 represent sequences used in the exemplification
CC of the present invention

Sequence 95 AA;

Query Match 92.9%; Score 26; DB 3; Length 95;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 66 VRDQLASSL 74

RESULT 15
ADU72822
ID ADU72822 standard; protein; 95 AA.
XX AC ADU72822;
XX 10-FEB-2005 (first entry)
DT

XX DE Signal peptide-containing polypeptide fragment, SEQ ID NO:1419.
XX KW Protein secretion; recombinant protein; diagnosis; mapping; forensic;
XX KW gene therapy; signal peptide.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1. .84
XX FT /label= Signal_peptide
XX FT Protein 85. .95
XX FT /label= Mature_polypeptide_fragment
XX FT
XX PN US6822072-B1.
XX PD 23-NOV-2004.
XX PF 21-DEC-1999; 99US-00471276.
XX PR 09-APR-1998; 98US-00057719.
XX PR 28-APR-1998; 98US-00069047.
XX PR 09-APR-1999; 99WO-IB000712.
XX PA (GEST) GENSET SA.
XX PI Edwards JDM, Duclert A, Giordano J;
XX WPI; 2004-812112/80.
XX N-PSDB; ADU72034.
XX
XX New expressed sequence tags and encoded human proteins useful for
XX diagnosing, preventing or treating diseases such as autoimmune disorders,
XX inflammation, wounds or infections, or in forensic or chromosome mapping
XX procedures.
XX Example 15; SEQ ID NO 1419; 72pp; English.
XX
XX The invention relates to an isolated or purified signal peptide
XX consisting of residues 1-16 of ADU72234 (signal peptide given separately
XX as ADU73026) which directs the extracellular secretion of a polypeptide
XX to which it is operably linked. The invention also relates to a method of
XX producing the signal peptide. The invention further discloses: isolated,
XX purified or enriched 5' expressed sequence tags (ESTs), many of which
XX encode all or a part of a secretory signal peptide; polypeptides encoded
XX by these ESTs (EST-related polypeptides); antibodies which recognize the
XX EST-related polypeptides; vectors and host cells comprising EST-related
XX nucleic acids of the invention; an array of ESTs; methods involving the
XX use of signal peptides of the invention to target polypeptides; and
XX methods involving the use of ESTs of the invention, for example, in
XX identifying a promoter in genomic DNA. The EST-encoded signal peptides of
XX the invention are useful for directing the secretion or import of a
XX recombinant polypeptide via the generation of protein fusions comprising
XX such signal peptides. The ESTs, EST-related polypeptides and methods of
XX the invention can be used for forensic procedures, chromosome mapping,
XX diagnostics, and therapeutic procedures, including gene therapy.
XX Sequences ADU72215-ADU72919 represent incomplete polypeptides comprising
XX a signal peptide which are encoded by the 5' ESTs shown in ADU71427-
XX ADU72131. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from the US patent office at
XX seqdata.uspto.gov/sequence.html?DocID=US6822072.
XX
XX Sequence 95 AA;

Query Match 92.9%; Score 26; DB 8; Length 95;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 66 VRDQLASSL 74

RESULT 16
ADZ73813
ID ADZ73813 standard; protein; 95 AA.
XX
AC ADZ73813;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human incomplete polypeptide including a signal peptide SEQ ID NO:1419.
XX
KW expressed sequence tag; EST; expression; protein secretion; diagnostic;
KW forensic; gene therapy; haplotype mapping.
XX
OS Homo sapiens.
XX
PN US2005106595-A1.
XX
PD 19-MAY-2005.
XX
PF 25-AUG-2004; 2004US-00926683.
XX
PR 09-APR-1998; 98US-00057719.
PR 28-APR-1998; 98US-00069047.
PR 09-APR-1999; 99WO-IB000712.
PR 21-DEC-1999; 99US-00471276.
XX
PA (GEST) GENSET SA.
XX
PI Dumas MEJ, Duclert A, Giordano J;
XX
XX WPI; 2005-384300/39.
XX N-PSDB; ADZ73025.
XX
XX New purified nucleic acid expressing secreted proteins useful in
XX forensic, gene therapy, and chromosome mapping procedures, and diagnosing
XX or treating cancer, atherosclerosis and autoimmune diseases, diabetes,
XX asthma and infections.
XX
XX Claim 1; SEQ ID NO 1419; 79pp; English.
XX
XX The invention relates to a novel purified nucleic acid (I) comprising any
XX of (ADZ72418-ADZ73205) or (ADZ73994-ADZ74016) and their complements; at
XX least 15 consecutive nucleotides of (I) and their complements; or any of
XX 788 nucleotide sequences encoding fully defined sequences of 16-255 amino
XX acids (ADZ73206-ADZ73993). The invention discloses 5' EST's derived from
XX mRNAs encoding secreted proteins. The 5' EST's may be used to obtain
XX cDNAs and genomic DNAs corresponding to the 5' ESTs. The methods and
XX compositions of the present invention are useful for expressing secreted
XX proteins or its portions (claimed) or to obtain antibodies capable of
XX specifically binding to the secreted proteins, and in diagnostic,
XX forensic, gene therapy, and chromosome mapping procedures, and for
XX designing expression vectors and secretion vectors. The present sequence
XX is used in the exemplification of the invention. Note: The sequence data
XX for this patent is not represented in the printed specification but is
XX based on sequence information supplied in electronic format from the
XX USPTO web site seqdata.uspto.gov/sequence.html; Document ID: 20050106595.
XX
XX Sequence 95 AA;

Query Match 92.9%; Score 26; DB 9; Length 95;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 66 VRDQLASSL 74

RESULT 17
ABU48350
ID ABU48350 standard; protein; 357 AA.
XX

AC ABU48350;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #33877.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Salmomella typhi.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA52220.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 76274; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 357 AA;

Query Match 92.9%; Score 26; DB 6; Length 357;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 VRDQIXXXL 9
| | | | | |
Db 255 VRDQIAKYL 263

RESULT 18
ADI67225
ID ADI67225 standard; protein; 465 AA.
XX
AC ADI67225;
XX
DT 24-FEB-2005 (revised)
DT 22-APR-2004 (first entry)
XX
DE Lactobacillus rhamnosus polypeptide sequence #99.
XX
KW Lactic acid bacteria; vaccine; bacterial infection; microbe;
KW milk-derived product; food product; food additive;
KW nutritional supplement; bioactive substance; probiotic supplement;
KW flavour; aroma; texture; nutritional value; food; microorganism;
KW antibacterial.
XX
OS Lactobacillus rhamnosus; strain HN001.
XX
PN US2004009490-A1.
XX
PD 15-JAN-2004.
XX
PF 03-OCT-2002; 2002US-00264213.
XX
PR 09-AUG-1999; 99US-0147852P.
PR 09-AUG-1999; 99US-0147853P.
PR 01-SEP-1999; 99US-0152031P.
PR 01-SEP-1999; 99US-0152032P.
PR 08-AUG-2000; 2000US-00634238.
PR 02-OCT-2001; 2001US-00971536.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Glenn M, Havukkala IJ, Lubbers M, Dekker J;
XX
DR WPI; 2004-090459/09.
DR N-PSDB; ADI67100.

New polynucleotide from Lactobacillus rhamnosus HN001 strain, useful for preparing a vaccine against bacterial infections or for modifying the flavor, aroma or nutritional benefits of a bioactive or probiotic supplement product.

Claim 21; SEQ ID NO 220; 54pp; English.

The present invention relates to the isolation of novel lactic acid bacteria (Lactobacillus rhamnosus strain HN001) polynucleotide sequences, and the proteins encoded by them. Also disclosed are oligonucleotide probes and primers, and genetic constructs comprising the polynucleotide sequences of the invention. The polynucleotide sequences are useful for preparing a vaccine against bacterial infections or for improving the properties of microbes used in the manufacture of milk-derived products, food products, food additives, nutritional supplements, bioactive substances or probiotic supplements, and for modifying the flavour, aroma, texture and/or nutritional value of foods. They are also useful for identifying microorganisms having a trait associated with the polynucleotide. The present sequence represents a novel L. rhamnosus polypeptide sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification. The complete sequence data for this patent was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov.

Sequence 465 AA;

Query Match 92.9%; Score 26; DB 8; Length 465;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;

DR WPI; 2005-028716/03.
XX
PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
PS Example 2; SEQ ID NO 16112; 25pp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16925 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 2559 AA;

Query Match 92.9%; Score 26; DB 9; Length 2559;
Best Local Similarity 66.7%; Pred. No. 5.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXXL 9
Db 268 VRDQIRSVL 276

RESULT 22
AAB16977
ID AAB16977 standard; peptide; 9 AA.
XX
AC AAB16977;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:33.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 207; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 9 AA;

Query Match 89.3%; Score 25; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXXL 9
Db 1 VRDQIXXXXL 9

RESULT 23
ABB72863
ID ABB72863 standard; peptide; 9 AA.
XX
AC ABB72863;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:33.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX Sequence 9 AA;

Query Match 89.3%; Score 25; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 1 VRDQIXXXL 9

RESULT 24

ABG71745
ID ABG71745 standard; peptide; 9 AA.

XX ABG71745;

DT 20-JAN-2003 (first entry)

XX TPO receptor, MPL, agonist peptide consensus sequence #1.

XX Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
KW complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
KW thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
KW T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
KW proliferation; growth; differentiation; haematopoietic cell;
KW platelet progenitor cell; immune disorder; thrombocytopenia;
KW disseminated intravascular coagulation; stem cell; transplantation;
KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.

XX Synthetic.

OS Key Location/Qualifiers

FH Key Location/Qualifiers
FT Misc-difference 6. .8
FT /label= Unknown
FT /note= "Residues are not defined within the
FT specification"

XX WO200278612-A2.

PD 10-OCT-2002.

XX 02-APR-2002; 2002WO-US010301.

XX 02-APR-2001; 2001US-0281183P.

XX (PURD) PURDUE PHARMA LP.

PI Soltis DA, Burch RM, Ogert RA;

XX WPI; 2003-040615/03.

XX New thrombopoietin synthebodies, useful for stimulating proliferation,
PT growth, or differentiation of hematopoietic cells, for treating or
PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.

XX Disclosure; Page 11; 97pp; English.

XX The invention discloses a variant of an immunoglobulin (Ig) variable
PS heavy or light chain domain that comprises at least one complementarity
CC determining region (CDR) and framework regions flanking the CDR. The CDR
CC also has added or substituted to it, at least one binding sequence which
CC is heterologous to the CDR and is an antigenic, agonistic sequence from a
CC thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
CC can be a binding sequence heterologous to the CDR, a cytotoxic T-
CC lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
CC cell sequence or a combination of each. The variant or thrombopoietin
CC synthebody, pharmaceutical and vaccine compositions are useful for
CC stimulating proliferation, growth or differentiation of haematopoietic
CC cells, particularly platelet progenitor cells. The variants are also
CC useful for treating or preventing haematopoietic or immune disorders
CC resulting from chemotherapy, radiation therapy, or bone marrow
CC transfusions (e.g. thrombocytopenia or disseminated intravascular
CC coagulation). Compositions comprising the synthebodies can be used for
CC the mobilisation, amplification and ex vivo expansion of stem cells and
CC committed precursor cells for autologous and allogeneic transplantation
CC as well as for the expansion of stem cells for gene therapy. They are
CC also useful as diagnostic or analytical reagents for studying the
CC function of thrombopoietin and its receptor in vivo or in vitro. The
CC sequence presented is the TPO receptor (MPL) agonist peptide consensus
CC sequence #1

XX Sequence 9 AA;

Query Match 89.3%; Score 25; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 1 VRDQIXXXL 9

RESULT 25

ADJ73014
ID ADJ73014 standard; peptide; 9 AA.

XX ADJ73014;

DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 468.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX Synthetic.

XX WO2003084477-A2.

PD 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX

DR WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 468; 97pp; English.

PS This invention relates to novel mammalian CDR mimetibodies, specific

XX portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is a TPO mimetic peptide sequence used to make a

CC mimetibody of the invention.

XX Sequence 9 AA;

SQ Query Match 89.3%; Score 25; DB 7; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 1 VRDQIXXXL 9

RESULT 26

ADJ52649

ID ADJ52649 standard; peptide; 9 AA.

XX ADJ52649;

XX 06-MAY-2004 (first entry)

DT CH1 deleted mimetibody-related peptide SeqID468.

DE

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

OS

XX Key Location/Qualifiers

FH Misc-difference 1. .9

FT /note= "All Xaa's in this sequence are unidentified amino

FT acids"

FT

XX WO2004002417-A2.

PN

XX 08-JAN-2004.

PD

XX 27-JUN-2003; 2003WO-US020347.

PF

XX 28-JUN-2002; 2002US-0392431P.

PR

XX (CEN2) CENTOCOR INC.

PA

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082870/08.

DR New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX Claim 2; SEQ ID NO 468; 129pp; English.

PS This invention relates to CH1 deleted mimetibodies (and the DNA sequences

XX which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an immunosuppressive,

CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,

CC antibacterial, virucide or fungicide activity. In addition, the disclosed

CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody

CC is useful for diagnosing or treating a disease condition in a cell,

CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart

CC failure), or neurodegenerative (for example multiple sclerosis, dementia

CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or

CC fungal infection). The present sequence is that of a peptide which may be

CC used during the creation of a mimetibody of the invention.

XX Sequence 9 AA;

SQ Query Match 89.3%; Score 25; DB 8; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 1 VRDQIXXXL 9

RESULT 27

ADJ51610

ID ADJ51610 standard; peptide; 9 AA.

XX ADJ51610;

XX 06-MAY-2004 (first entry)

DT CH1 deleted mimetibody-related peptide SeqID468.

DE

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunological disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

OS

XX Key Location/Qualifiers

FH Misc-difference 1. .9

FT /note= "All Xaa's in this sequence are unidentified amino

FT acids"

FT

XX WO2004002424-A2.

PN

XX 08-JAN-2004.
PD
XX
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
XX
PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 468; 123pp; English.
XX
CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 9 AA;

Query Match 89.3%; Score 25; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db |||||
1 VRDQIXXXL 9

RESULT 28
AAY22357
ID AAY22357 standard; peptide; 10 AA.
XX
AC AAY22357;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 8.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
AC AAY22357;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 8.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX

PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db |||||
2 VRDQVSWAL 10

RESULT 29
AAY22360
ID AAY22360 standard; peptide; 10 AA.
XX
AC AAY22360;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 11.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX

PS Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding peptide of the invention. The peptide has: (i) a molecular weight below 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10 mM. The peptides are used to treat conditions requiring a thrombopoietin agonist, particularly haematological disorders or thrombocytopaenia, especially resulting from chemotherapy, radiation therapy or bone marrow transfusions. Also when labelled they may be used for diagnosis (detecting TR on cells); for studying mechanisms of haematopoiesis; for in vitro expansion of megakaryocytes and committed progenitor cells, and for the development/identification of other TR agonists. The compounds accelerate thrombocyte regeneration

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 2 VRDQVSWAL 10

RESULT 30
AAB16981

ID AAB16981 standard; peptide; 10 AA.

XX AAB16981;

AC AAB16981;

XX 31-OCT-2000 (first entry)

DT TPO-mimetic peptide sequence SEQ ID NO:37.

DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer; autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF; immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1; cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor; vascular endothelial growth factor; matrix metalloproteinase; asthma; thrombosis; pharmaceutical.

XX Synthetic.

OS WO200024782-A2.

XX 04-MAY-2000.

PD 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

DR Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

PT Claim 19; Page 208; 608pp; English.

PS The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 3; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 2 VRDQVSWAL 10

RESULT 31
ABB72867

ID ABB72867 standard; peptide; 10 AA.

XX ABB72867;

AC ABB72867;

XX 05-APR-2002 (first entry)

DT TPO mimetic peptide SEQ ID NO:37.

DE Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP; TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianaemic; anorectic; antiinfertility; haemostatic; dermatological; neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

PN 08-NOV-2001.

PD 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

PR (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

PI WPI; 2002-130313/17.

DR Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

PS The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 5; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 2 VRDQVSWAL 10

RESULT 32
ADJ73018
ID ADJ73018 standard; peptide; 10 AA.

AC ADJ73018;

DT 06-MAY-2004 (first entry)

DE TPO mimetic peptide sequence SeqID 472.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX Synthetic.

PN WO2003084477-A2.

XX 16-OCT-2003.

PF 24-MAR-2003; 2003WO-US009139.

PR 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX WPI; 2003-804237/75.

DR New CDR mimetibody comprising a portion of a heavy or light chain
XX variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 472; 97pp; English.

PS This invention relates to novel mammalian CDR mimetibodies, specific
XX portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 7; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 2 VRDQVSWAL 10

RESULT 33
ADJ52653

ID ADJ52653 standard; peptide; 10 AA.

AC ADJ52653;

XX 06-MAY-2004 (first entry)

DE CH1 deleted mimetibody-related peptide SeqID472.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

XX WO2004002417-A2.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US020347.

PR 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082870/08.

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

PS Claim 2; SEQ ID NO 472; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 8; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 2 VRDQVSWAL 10
| | | | : |

RESULT 34
ADJ51614
ID ADJ51614 standard; peptide; 10 AA.
XX
AC ADJ51614;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID472.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiinflammatory; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutolowski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 472; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiinflammatory, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 8; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 2 VRDQVSWAL 10
| | | | : |

RESULT 35
AAB16989
ID AAB16989 standard; peptide; 11 AA.
XX
AC AAB16989;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:45.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
PI
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 211; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC p3, and p4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX Sequence 11 AA;

Query Match 89.3%; Score 25; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 3 VRDQIXXXL 11
|||||

RESULT 36
ABB72875
ID ABB72875 standard; peptide; 11 AA.

XX ABB72875;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:45.

DE Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS Synthetic.

PN WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX Sequence 11 AA;

Query Match 89.3%; Score 25; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 3 VRDQIXXXL 11
|||||

RESULT 37

ADJ52661
ID ADJ52661 standard; peptide; 11 AA.

XX ADJ52661;

DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID480.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

OS Unidentified.
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1..9
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX WO2004002417-A2.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US020347.

XX 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;

XX WPI; 2004-082870/08.

XX

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 480; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 11 AA;

Query Match 89.3%; Score 25; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 3 VRDQIXXXL 11

RESULT 38
AAY22388
ID AAY22388 standard; peptide; 15 AA.
XX
AC AAY22388;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 39.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwiria SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwiria SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10

CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;

Query Match 89.3%; Score 25; DB 2; Length 15;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 6 VRDQVLAFL 14

RESULT 39
AAY22375
ID AAY22375 standard; peptide; 15 AA.
XX
AC AAY22375;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 26.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwiria SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX

CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;

Query Match 89.3%; Score 25; DB 2; Length 15;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 6 VRDQVTAWL 14

RESULT 40
AAY22394
ID AAY22394 standard; peptide; 15 AA.
XX
AC AAY22394;
XX 27-SEP-1999 (first entry)
DT TPO receptor binding peptide sequence, SEQ ID NO. 45.
DE TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
XX haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX US5932546-A.
PN 03-AUG-1999.
XX 04-OCT-1996; 96US-00726464.
PF 04-OCT-1996; 96US-00726464.
XX 04-OCT-1996; 96US-00726464.
PR (GLAX) GLAXO WELLCOME INC.
XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX WPI; 1999-457122/38.
DR New low molecular weight thrombopoietin agonists, particularly peptides,
XX for treatment of hematological disease and thrombocytopenia.
XX Disclosure; Col 13-14; 36pp; English.
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;

QY 1 VRDQIXXXL 9
DB 6 VRDQVIARL 14

RESULT 41
AAY22383
ID AAY22383 standard; peptide; 15 AA.
XX
AC AAY22383;
XX 27-SEP-1999 (first entry)
DT
XX

QY 1 VRDQIXXXL 9
DB 6 VRDQVAMYL 14

RESULT 42
ABO67958
ID ABO67958 standard; protein; 219 AA.
XX
AC ABO67958;
XX 29-JUL-2004 (first entry)
DT Pseudomonas aeruginosa polypeptide #133.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
OS US6551795-B1.
PN 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
PF 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
PR

Query Match 89.3%; Score 25; DB 2; Length 15;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DE TPO receptor binding peptide sequence, SEQ ID NO. 34.
XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX US5932546-A.
PN 03-AUG-1999.
XX 04-OCT-1996; 96US-00726464.
PF 04-OCT-1996; 96US-00726464.
XX 04-OCT-1996; 96US-00726464.
PR (GLAX) GLAXO WELLCOME INC.
XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX WPI; 1999-457122/38.
DR New low molecular weight thrombopoietin agonists, particularly peptides,
XX for treatment of hematological disease and thrombocytopenia.
XX Disclosure; Col 13-14; 36pp; English.
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;

QY 1 VRDQIXXXL 9
DB 6 VRDQVAMYL 14

RESULT 42
ABO67958
ID ABO67958 standard; protein; 219 AA.
XX
AC ABO67958;
XX 29-JUL-2004 (first entry)
DT Pseudomonas aeruginosa polypeptide #133.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
OS US6551795-B1.
PN 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
PF 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
PR

Query Match 89.3%; Score 25; DB 2; Length 15;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

XX (GENO-) GENOME THERAPEUTICS CORP.
PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD01529.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 16704; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX Sequence 219 AA;
SQ

Query Match 89.3%; Score 25; DB 7; Length 219;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 113 VRDQVRGAL 121

RESULT 43
ABB61401
ID ABB61401 standard; protein; 250 AA.
XX ABB61401;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 10995.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW Drosophila melanogaster.
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL05504.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 10995; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 250 AA;
SQ

Query Match 89.3%; Score 25; DB 4; Length 250;
Best Local Similarity 55.6%; Pred. No. 7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 220 IRDQIAERL 228

RESULT 44
AAY43936
ID AAY43936 standard; protein; 276 AA.
XX AAY43936;
AC

XX 21-DEC-1999 (first entry)
DT
XX Human protein kinase #6.
DE
XX Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.
XX Homo sapiens.
OS
XX US5958784-A.
PN
XX 28-SEP-1999.
PD
XX 25-MAR-1992; 92US-00857224.
PF
XX 25-MAR-1992; 92US-00857224.
PR
XX (BENN/) BENNER S A.
PA Benner SA;
XX WPI; 1999-570766/48.
DR
XX Predicting the folded structure of proteins.
PT
XX Disclosure; Col 209-210; 113pp; English.
PS

Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and parsing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active site
CC assignments and co-variation analysis. The predicted secondary structures
CC are useful for identifying antigenic sites on a protein molecule, as
CC guides for site directed mutagenesis studies, and for understanding the

CC interaction of a protein with other molecules
XX
SQ Sequence 276 AA;

Query Match 89.3%; Score 25; DB 2; Length 276;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 82 VRDQVIAVL 90

RESULT 45
ADI57223
ID ADI57223 standard; protein; 314 AA.
XX
AC ADI57223;
XX
DT 22-APR-2004 (first entry)
XX
DE CDC7L1 dominant negative mutant amino acid sequence CDC7Sc.
XX
KW cell cycle arrest; CK2-specific siRNA molecule; short interfering RNA;
KW CK2 inhibition; PIM1-specific siRNA molecule; PIM1 inhibition;
KW Hbol-specific siRNA molecule; Hbol inhibition; cytostatic; gene therapy;
KW cancer; proliferative disorder; human; CDC7L1; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004007754-A2.
XX
PD 22-JAN-2004.
XX
PF 14-JUL-2003; 2003WO-US022164.
XX
PR 12-JUL-2002; 2002US-0395443P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Hitoshi Y, Jenkins Y, Markovtsov V;
XX
DR WPI; 2004-122975/12.
XX
PT Identifying a compound that modulates cell cycle arrest, useful for
PT developing therapeutic reagents for treating cancer comprising contacting
PT a cell comprising a target polypeptide with the compound.
XX
PS Disclosure; Fig 23; 180pp; English.
XX
CC The present invention describes a method for identifying a compound (C)
CC that modulates cell cycle arrest. The method comprises contacting a cell
CC comprising a target polypeptide with the compound (C), where the target
CC polypeptide encoded by the complement of a nucleic acid that hybridises
CC under stringent conditions to a nucleic acid encoding a polypeptide
CC having an amino acid sequence selected from 18 148-1408 amino acid
CC sequences (SEQ ID NO: 2-36, even numbers only). Also described: (1)
CC modulating cell cycle arrest in a subject; (2) a CK2-specific short
CC interfering RNA (siRNA) molecule comprising the sequence: (I)
CC AACATTGAATTAGATCCACGT, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; (3) inhibiting expression of a CK2 gene in a cell;
CC (4) a PIM1-specific siRNA molecule comprising the sequence: (II)
CC AAAACTCCGAGTGAAGTGGTC, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; (5) inhibiting expression of a PIM1 gene in a cell;
CC (6) an Hbol-specific siRNA molecule comprising the sequence: (III)
CC AACTGAGCAAGTGGTTGATTT, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; and (7) inhibiting expression of an Hbol gene in a
CC cell. (C) has cytostatic activity, and can be used in gene therapy. The
CC method is useful for identifying a compound (c) that modulates cell cycle
CC arrest. Compounds that modulate cell cycle arrest are useful for
CC developing therapeutic reagents for treating cancer and other
CC proliferative disorders. The present sequence represents a dominant ,

CC negative mutant of human CDC7L1, which is used in the exemplification of
CC the present invention.
XX
SQ Sequence 314 AA;

Query Match 89.3%; Score 25; DB 8; Length 314;
Best Local Similarity 55.6%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 112 VRDQVIAVL 120

RESULT 46
ADI57229
ID ADI57229 standard; protein; 314 AA.
XX
AC ADI57229;
XX
DT 22-APR-2004 (first entry)
XX
DE CDC7L1 dominant negative mutant amino acid sequence.
XX
KW cell cycle arrest; CK2-specific siRNA molecule; short interfering RNA;
KW CK2 inhibition; PIM1-specific siRNA molecule; PIM1 inhibition;
KW Hbol-specific siRNA molecule; Hbol inhibition; cytostatic; gene therapy;
KW cancer; proliferative disorder; human; CDC7L1; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004007754-A2.
XX
PD 22-JAN-2004.
XX
PF 14-JUL-2003; 2003WO-US022164.
XX
PR 12-JUL-2002; 2002US-0395443P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Hitoshi Y, Jenkins Y, Markovtsov V;
XX
DR WPI; 2004-122975/12.
XX
PT Identifying a compound that modulates cell cycle arrest, useful for
PT developing therapeutic reagents for treating cancer comprising contacting
PT a cell comprising a target polypeptide with the compound.
XX
PS Example 18; Fig 26; 180pp; English.
XX
CC The present invention describes a method for identifying a compound (C)
CC that modulates cell cycle arrest. The method comprises contacting a cell
CC comprising a target polypeptide with the compound (C), where the target
CC polypeptide encoded by the complement of a nucleic acid that hybridises
CC under stringent conditions to a nucleic acid encoding a polypeptide
CC having an amino acid sequence selected from 18 148-1408 amino acid
CC sequences (SEQ ID NO: 2-36, even numbers only). Also described: (1)
CC modulating cell cycle arrest in a subject; (2) a CK2-specific short
CC interfering RNA (siRNA) molecule comprising the sequence: (I)
CC AACATTGAATTAGATCCACGT, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; (3) inhibiting expression of a CK2 gene in a cell;
CC (4) a PIM1-specific siRNA molecule comprising the sequence: (II)
CC AAAACTCCGAGTGAAGTGGTC, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; (5) inhibiting expression of a PIM1 gene in a cell;
CC (6) an Hbol-specific siRNA molecule comprising the sequence: (III)
CC AACTGAGCAAGTGGTTGATTT, where the siRNA molecule is from 21-30 nucleotide
CC base pairs in length; and (7) inhibiting expression of an Hbol gene in a
CC cell. (C) has cytostatic activity, and can be used in gene therapy. The
CC method is useful for identifying a compound (c) that modulates cell cycle
CC arrest. Compounds that modulate cell cycle arrest are useful for
CC developing therapeutic reagents for treating cancer and other

CC proliferative disorders. The present sequence represents a dominant
CC negative mutant of human CDC7L1, which is used in the exemplification of
CC the present invention.
XX
SQ Sequence 314 AA;

Query Match 89.3%; Score 25; DB 8; Length 314;
Best Local Similarity 55.6%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 112 VRDQVIAVL 120

RESULT 47
ADN18947
ID ADN18947 standard; protein; 507 AA.
XX
AC ADN18947;
XX
DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #1600.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 1600; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 507 AA;

Query Match 89.3%; Score 25; DB 8; Length 507;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 112 VRDQVIAVL 120

RESULT 48
ADS42509
ID ADS42509 standard; protein; 519 AA.
XX
AC ADS42509;
XX
DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #20939.
XX
DE Bacterial polypeptide #20939.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 20939; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 519 AA;

Query Match 89.3%; Score 25; DB 8; Length 519;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||||| :
Db 82 VRDQIATEI 90

RESULT 49
ABU22938
ID ABU22938 standard; protein; 530 AA.

AC ABU22938;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #8465.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

KW Bordetella pertussis.

OS WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

PI WPI; 2003-029926/02.

DR N-PSDB; ACA26808.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 50862; 1766pp; English.

PS The invention relates to an isolated nucleic acid comprising any one of

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 530 AA;

Query Match 89.3%; Score 25; DB 6; Length 530;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||||| :
Db 41 VRDQVAQGL 49

RESULT 50
ABO76610

ID ABO76610 standard; protein; 564 AA.

XX ABO76610;

AC 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #8785.

DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

KW Pseudomonas aeruginosa.

XX OS US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

PI WPI; 2003-615309/58.

XX N-PSDB; ABD10181.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.

XX (1) a vector comprising a promoter operably linked to the nucleic acid.

PS Disclosure; SEQ ID NO 25356; 455pp; English.

XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and

CC therapy of pathological conditions, as molecular targets for diagnostics,

CC prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a P. aeruginosa nucleic acid, as components of

CC effective antibacterial targets, as targets for antibacterial drugs,

CC including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused

CC infection, and in detection of P. aeruginosa sequences or other sequences

CC of Pseudomonas species using biochip technology. Sequences ABO67826-

CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format from USPTO at

CC seqdata.uspto.gov/sequence.html

XX

SQ Sequence 564 AA;

Query Match 89.3%; Score 25; DB 7; Length 564;

Best Local Similarity 55.6%; Pred. No. 1.8e+03;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXL 9

DB 234 VRDQWADL 242

Search completed: May 12, 2006, 10:37:49

Job time : 94.4846 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 10.9231 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	27	96.4	543	2 T02022	hypothetical prote
2	27	96.4	1113	2 T47381	hypothetical prote
3	27	96.4	1198	2 E86402	hypothetical prote
4	27	96.4	1201	2 F86386	hypothetical prote
5	27	96.4	1305	2 H96559	hypothetical prote
6	27	96.4	1472	2 A84470	hypothetical prote
7	27	96.4	1756	2 T02599	hypothetical prote
8	26	92.9	211	2 T01194	hypothetical prote
9	26	92.9	357	2 AD1062	protein kinase [im
10	26	92.9	1236	2 B36329	hypothetical prote
11	26	92.9	1312	2 D85066	hypothetical prote
12	25	89.3	115	2 S09866	hypothetical prote
13	25	89.3	193	2 D83620	hypothetical prote
14	25	89.3	243	2 T23381	hypothetical prote
15	25	89.3	327	2 S14268	peroxidase (EC 1.1
16	25	89.3	507	2 A25228	protein kinase CDC
17	25	89.3	610	2 F82192	ABC transporter, A
18	25	89.3	621	2 AF3016	Na+/H+ antiporter
19	25	89.3	642	2 B98268	probable sodium/hy
20	25	89.3	686	2 F89895	ATP-dependent DNA
21	25	89.3	780	2 T00122	hypothetical 88.9K
22	25	89.3	1430	2 T21910	hypothetical prote
23	25	89.3	3871	2 T22812	hypothetical prote
24	24	85.7	103	2 G82442	Su1l family protei
25	24	85.7	109	2 A87425	hypothetical prote
26	24	85.7	166	2 T39825	hypothetical prote
27	24	85.7	177	2 E83662	hypothetical prote
28	24	85.7	185	2 S66113	hypothetical prote
29	24	85.7	193	1 S39401	hydrogenase matura

30	24	85.7	205	2 B82601	transcription regu
31	24	85.7	235	2 F81658	type III secretion
32	24	85.7	250	2 E70104	1-acylglycerol-3-p
33	24	85.7	257	2 AE0127	Conserved hypothet
34	24	85.7	259	2 A69965	ketoacyl reductase
35	24	85.7	260	2 AG0861	probable RNA pseud
36	24	85.7	260	2 C91085	hypothetical prote
37	24	85.7	260	2 E85930	hypothetical prote
38	24	85.7	260	2 C65061	hypothetical prote
39	24	85.7	302	2 T22929	hypothetical prote
40	24	85.7	303	1 CPCYB	carboxypeptidase B
41	24	85.7	315	2 H87447	conserved hypothet
42	24	85.7	324	2 C87271	general secretion
43	24	85.7	351	2 T29369	hypothetical prote
44	24	85.7	369	2 T04917	hypothetical prote
45	24	85.7	376	2 S45107	hypothetical prote
46	24	85.7	448	2 T32710	hypothetical prote
47	24	85.7	471	2 S61029	hypothetical prote
48	24	85.7	505	2 S64837	hypothetical prote
49	24	85.7	512	1 PQ0646	ferredoxin-nitrite
50	24	85.7	568	2 G83558	hypothetical prote
51	24	85.7	600	2 H81331	excinuclease ABC c
52	24	85.7	600	2 H84176	medium-chain acyl-
53	24	85.7	628	2 S77374	cell division prot
54	24	85.7	628	2 AI0382	peptidylprolyl iso
55	24	85.7	648	2 C83023	probable oxidoredu
56	24	85.7	654	1 HHBYS1	dnAK-type molecula
57	24	85.7	781	2 AI3014	methyl-accepting c
58	24	85.7	788	2 F98269	hypothetical prote
59	24	85.7	891	2 S75940	methyl-accepting c
60	24	85.7	997	2 T28872	hypothetical prote
61	24	85.7	1114	2 T49517	p63 related protei
62	24	85.7	1175	2 T46124	hypothetical prote
63	24	85.7	1175	2 D85089	hypothetical prote
64	24	85.7	1203	2 T04294	hypothetical prote
65	24	85.7	1489	2 G71406	probable retroviru
66	24	85.7	1633	2 T01879	hypothetical prote
67	23	82.1	83	2 S28137	gas vesicle protei
68	23	82.1	91	2 S78225	heat shock protein
69	23	82.1	140	2 E86437	protein F28K20.11
70	23	82.1	149	2 AF0127	conserved hypothet
71	23	82.1	154	2 E72708	hypothetical prote
72	23	82.1	171	2 A86578	CT676 hypothetical
73	23	82.1	171	2 D72045	conserved hypothet
74	23	82.1	173	2 C71484	hypothetical prote
75	23	82.1	177	2 F81746	conserved hypothet
76	23	82.1	180	2 D82550	conserved hypothet
77	23	82.1	203	2 AC1595	hypothetical prote
78	23	82.1	236	2 T41412	probable hydantoin
79	23	82.1	248	2 S30403	hypothetical prote
80	23	82.1	265	2 B70904	hypothetical prote
81	23	82.1	303	2 H97122	hypothetical prote
82	23	82.1	303	2 A48600	probable sn2-acylg
83	23	82.1	308	2 G84210	ABC transporter, A
84	23	82.1	317	2 AI3567	tetratricopeptide
85	23	82.1	319	2 S62191	5-methyltetrahydro
86	23	82.1	328	2 A84167	UDP-glucose 4-epim
87	23	82.1	331	1 A70032	conserved hypothet
88	23	82.1	333	1 A38094	D-lactate dehydrog
89	23	82.1	334	2 A83662	transcription regu
90	23	82.1	337	1 S29296	D-lactate dehydrog
91	23	82.1	349	2 AE2369	hypothetical prote
92	23	82.1	360	2 G82551	phage-related prot
93	23	82.1	382	2 S32148	exo-alpha-sialidas
94	23	82.1	383	2 S34810	mating factor MAT1
95	23	82.1	389	2 E65182	bacteriophage N4 a
96	23	82.1	390	2 A86065	UDP-N-acetyl gluco
97	23	82.1	390	2 G91218	UDP-N-acetyl gluco
98	23	82.1	392	2 B75466	sensor histidine k
99	23	82.1	403	2 JC7708	tocopherol associa
100	23	82.1	405	2 A82130	transcription regu
101	23	82.1	409	1 JN0885	N-carbamyl-L-amino
102	23	82.1	409	2 D84488	probable Tail-like

103 23 82.1 426 2 H71636 3-oxoacyl-[acyl-ca
104 23 82.1 426 2 D87457 3-oxoacyl-(acyl-ca
105 23 82.1 427 2 A84155 hypothetical prote
106 23 82.1 428 2 B97848 hypothetical prote
107 23 82.1 433 2 A84619 probable serine ca
108 23 82.1 446 2 F82609 aminopeptidase P X
109 23 82.1 451 2 F87659 conserved hypothet
110 23 82.1 460 2 I57546 Rabin3 - rat
111 23 82.1 464 2 T24011 hypothetical prote
112 23 82.1 470 2 C70641 hypothetical prote
113 23 82.1 495 2 T52515 related to n-alkan
114 23 82.1 503 1 A35376 cholesterol 7alpha
115 23 82.1 504 2 S39399 cholesterol 7alpha
116 23 82.1 504 2 G84251 hypothetical prote
117 23 82.1 527 2 D70595 probable ATP-depen
118 23 82.1 536 2 E82559 virulence factor X
119 23 82.1 544 2 E87010 probable ATP-depen
120 23 82.1 557 2 F89901 conserved hypoteht
121 23 82.1 566 2 B28391 alpha-amylase (EC
122 23 82.1 566 2 S14063 alpha-amylase (EC
123 23 82.1 566 2 C70482 repair excision nu
124 23 82.1 575 1 VCLJHD env polypeptin pr
125 23 82.1 575 2 S44018 aerobactin biosynt
126 23 82.1 579 2 F69157 excinuclease ABC c
127 23 82.1 588 1 BVECUC excinuclease ABC,
128 23 82.1 588 2 D85808 excinuclease ABC,
129 23 82.1 588 2 C90960 excinuclease ABC s
130 23 82.1 592 2 D95393 hypothetical prote
131 23 82.1 593 2 T44479 conserved hypothet
132 23 82.1 594 2 S09857 hypothetical prote
133 23 82.1 598 2 AG0890 probable arylsulfa
134 23 82.1 610 2 AD0749 excinuclease ABC c
135 23 82.1 610 2 AG0227 excinuclease ABC c
136 23 82.1 610 2 S77337 ABC-type transport
137 23 82.1 611 2 T44880 probable acyl-coA
138 23 82.1 611 2 H70938 probable fadE5 pro
139 23 82.1 613 1 S74461 ABC transporter sl
140 23 82.1 620 2 A83182 hypothetical prote
141 23 82.1 626 1 S77286 excinuclease ABC c
142 23 82.1 627 2 A12146 excinuclease ABC c
143 23 82.1 669 2 D82739 excinuclease ABC s
144 23 82.1 674 2 T20571 hypothetical prote
145 23 82.1 679 2 T20568 hypothetical prote
146 23 82.1 693 2 G82618 pilus biogenesis p
147 23 82.1 695 2 T02630 hypothetical prote
148 23 82.1 702 2 T52634 nematode responsiv
149 23 82.1 707 2 T32410 hypothetical prote
150 23 82.1 723 2 I39066 N-methyl-D-asparta
151 23 82.1 737 2 T46040 hypothetical prote
152 23 82.1 739 2 S47772 biotin sulfoxide r
153 23 82.1 739 2 D91183 biotin sulfoxide r
154 23 82.1 739 2 H86029 biotin sulfoxide r
155 23 82.1 822 2 T33163 hypothetical prote
156 23 82.1 888 2 T31131 pyruvate, phosphat
157 23 82.1 909 2 H86350 hypothetical prote
158 23 82.1 1008 2 T18832 probable RNA helic
159 23 82.1 1071 2 S38164 ATP-binding protei
160 23 82.1 1312 2 T30845 probable DNA repai
161 23 82.1 1314 2 T47331 hypothetical prote
162 23 82.1 1385 2 T21706 hypothetical prote
163 23 82.1 2004 2 AC0314 probable membrane
164 23 82.1 2535 2 AC0304 probable hemolysin
165 23 82.1 6420 2 T30283 polyketide synthas
166 22 78.6 60 2 AC2981 hypothetical prote
167 22 78.6 94 2 S69114 fibrinogen alpha c
168 22 78.6 99 2 S74615 hypothetical prote
169 22 78.6 100 2 T45169 flagellar hook-bas
170 22 78.6 103 2 G87367 hypothetical prote
171 22 78.6 116 2 D81945 probable pilus bio
172 22 78.6 116 2 G81159 probable type IV p
173 22 78.6 123 2 B45355 hypothetical prote
174 22 78.6 123 2 S64248 hypothetical prote
175 22 78.6 126 2 S74493 thioredoxin M-2 -

176 22 78.6 137 2 T20957 hypothetical prote
177 22 78.6 139 2 C69950 arsenate reductase
178 22 78.6 139 2 T33713 conserved hypothet
179 22 78.6 147 2 I46623 rearranged T-cell
180 22 78.6 151 2 AF0014 conserved hypothet
181 22 78.6 154 2 E82945 galactoside O-acet
182 22 78.6 160 2 G95113 regulatory protein
183 22 78.6 160 2 H97982 regulatory protein
184 22 78.6 169 2 C96494 protein F7F22.12 [
185 22 78.6 196 2 D82143 conserved hypothet
186 22 78.6 201 2 AC2371 hypothetical prote
187 22 78.6 204 1 MNZHS nonstructural prot
188 22 78.6 204 1 MNZSV nonstructural prot
189 22 78.6 208 2 A82077 MutT/nudix family
190 22 78.6 209 2 A71313 probable V-type AT
191 22 78.6 212 2 G82176 probable outer mem
192 22 78.6 213 2 T37039 hypothetical prote
193 22 78.6 217 2 S66155 probable DNA-direc
194 22 78.6 221 2 D87469 transcription regu
195 22 78.6 222 2 AD1397 ABC transporter, A
196 22 78.6 223 2 AG1772 ABC transporter, A
197 22 78.6 232 2 C95245 hypothetical prote
198 22 78.6 232 2 H98109 2,3,4,5-tetrahydro
199 22 78.6 232 2 A69131 conserved hypothet
200 22 78.6 236 2 A11558 tetrahydrodipicoli
201 22 78.6 236 2 F69866 tetrahydrodipicoli
202 22 78.6 236 2 AC1201 tetrahydrodipicoli
203 22 78.6 236 2 H71056 probable arylmalon
204 22 78.6 239 2 B83296 probable transcrip
205 22 78.6 239 2 T34733 hypothetical prote
206 22 78.6 240 2 S43983 tetrahydrodipicoli
207 22 78.6 245 2 S43565 R01H10.4 protein (
208 22 78.6 252 2 E86950 conserved hypothet
209 22 78.6 256 2 A86660 acetyltransferase
210 22 78.6 261 2 B83969 RNA polymerase spo
211 22 78.6 267 2 S15486 probable ABC-type
212 22 78.6 269 2 B71365 probable phosphome
213 22 78.6 270 2 D90113 hypothetical prote
214 22 78.6 276 2 F87679 chemotaxis protein
215 22 78.6 278 2 AD0146 probable ABC trans
216 22 78.6 280 2 T09576 Lbx1 transcription
217 22 78.6 281 1 D70570 probable acyl-CoA
218 22 78.6 288 2 E70580 hypothetical prote
219 22 78.6 293 2 AC2780 tolB protein [limpo
220 22 78.6 293 2 E97559 tolB protein precu
221 22 78.6 295 2 S23585 carboxyphosphonoen
222 22 78.6 297 2 G69525 formylmethanofuran
223 22 78.6 298 2 C69819 ABC transporter (A
224 22 78.6 300 2 AH3386 ATP-dependent prot
225 22 78.6 300 2 AE3328 hflC protein (EC 3
226 22 78.6 305 2 AF0305 sugar binding prot
227 22 78.6 306 2 AG2698 cation efflux syst
228 22 78.6 306 2 G97480 hypothetical prote
229 22 78.6 308 2 E98259 probable transcrip
230 22 78.6 308 2 AD3025 transcription regu
231 22 78.6 309 2 A42921 transcription elon
232 22 78.6 321 2 JS0772 37K glycoprotein -
233 22 78.6 332 2 F82140 C4-dicarboxylate-b
234 22 78.6 334 2 A10795 probable receptor/
235 22 78.6 339 2 G83900 hypothetical prote
236 22 78.6 344 2 D86362 hypothetical prote
237 22 78.6 349 2 T50372 probable ATP-depen
238 22 78.6 355 2 T36273 hypothetical prote
239 22 78.6 366 2 T16119 hypothetical prote
240 22 78.6 374 2 C87134 conserved hypothet
241 22 78.6 376 2 A75360 hypothetical prote
242 22 78.6 377 2 S77634 exopolysaccharide
243 22 78.6 379 2 D70846 probable metaA prot
244 22 78.6 381 2 C70711 hypothetical prote
245 22 78.6 381 2 AB2357 hypothetical prote
246 22 78.6 382 2 T45301 homoserine o-acety
247 22 78.6 386 2 T09019 phosphoprotein pho
248 22 78.6 386 2 A96532 hypothetical prote

249	22	78.6	388	2	C85323	protein phosphatas	322	22	78.6	875	2	F70755	hypothetical prote
250	22	78.6	389	2	F82809	conserved hypothet	323	22	78.6	883	2	A71434	probable RNA helic
251	22	78.6	389	2	G95939	probable hippurate	324	22	78.6	884	2	C70729	hypothetical prote
252	22	78.6	391	2	AG0533	probable drug effl	325	22	78.6	901	1	RGECMT	regulatory protein
253	22	78.6	398	2	JC2516	formaldehyde dismu	326	22	78.6	901	2	AG0996	Malt regulatory pr
254	22	78.6	398	2	S29819	serpin - barley	327	22	78.6	901	2	D91161	positive regulator
255	22	78.6	401	2	T15607	hypothetical prote	328	22	78.6	901	2	D86007	positive regulator
256	22	78.6	403	2	D82076	tRNA nucleotidyltr	329	22	78.6	903	2	AI0015	maltose regulon po
257	22	78.6	403	2	D96961	carboxyl-terminal	330	22	78.6	956	1	A46016	thrombospondin 3 -
258	22	78.6	409	2	B25156	cellulase (EC 3.2.	331	22	78.6	956	2	G70327	isoleucine-tRNA li
259	22	78.6	410	2	A45457	alpha 1-proteinase	332	22	78.6	956	2	A57121	thrombospondin 3 p
260	22	78.6	411	2	H83788	involved in spore	333	22	78.6	968	2	T46568	ATP-dependent RNA
261	22	78.6	416	2	B64132	tRNA adenyllytrans	334	22	78.6	1003	2	T24545	hypothetical prote
262	22	78.6	416	2	H95920	probable membrane-	335	22	78.6	1052	2	H83909	cell wall-associat
263	22	78.6	417	2	A55524	aspartic proteinas	336	22	78.6	1060	2	T30347	vitellogenin conve
264	22	78.6	418	2	S42072	aspartic proteinas	337	22	78.6	1078	2	T30879	dynein heavy chain
265	22	78.6	419	2	T10652	hypothetical prote	338	22	78.6	1087	2	D84767	probable pre-mRNA
266	22	78.6	420	2	T08691	hypothetical prote	339	22	78.6	1090	2	C86450	FSD14.27 protein -
267	22	78.6	424	2	T01383	GTPase-activating	340	22	78.6	1386	2	T49316	profilaggrin relat
268	22	78.6	429	2	AH1009	sun protein [impor	341	22	78.6	1599	2	T15854	hypothetical prote
269	22	78.6	431	2	AG1884	glycolate oxidase	342	22	78.6	1638	2	T30313	chemotaxis protein
270	22	78.6	436	2	T22253	hypothetical prote	343	22	78.6	1676	2	E71410	probable centromer
271	22	78.6	454	2	G81828	probable periplasm	344	22	78.6	1763	2	T17465	rifamycin polyketi
272	22	78.6	456	2	E83609	probable aminotran	345	22	78.6	2051	2	T13164	plexin B - fruit f
273	22	78.6	458	2	E69528	cobyrinic acid a,c	346	22	78.6	2116	2	A26655	myosin heavy chain
274	22	78.6	458	2	G83690	hypothetical prote	347	22	78.6	2166	2	G70163	hypothetical prote
275	22	78.6	459	2	S33424	repSA protein - St	348	22	78.6	2472	2	E83594	still frameshift p
276	22	78.6	466	2	AH3100	amidohydrolase [im	349	22	78.6	2475	2	T00047	gellan lyase (EC 4
277	22	78.6	466	2	B98186	probable hydrolase	350	22	78.6	2695	2	S54974	type 1 inositol 1,
278	22	78.6	468	2	D69716	involved in spore	351	22	78.6	2713	2	A55713	inositol 1,4,5-tri
279	22	78.6	473	2	H84550	probable obtusifol	352	22	78.6	2734	2	B36579	inositol 1,4,5-tri
280	22	78.6	475	2	B81047	secretion protein,	353	22	78.6	2749	1	ACMSIT	inositol 1,4,5-tri
281	22	78.6	484	1	B69056	IMP dehydrogenase	354	22	78.6	2749	2	A36579	inositol 1,4,5-tri
282	22	78.6	486	2	H82377	ABC transporter, A	355	22	78.6	3413	2	T17467	rifamycin polyketi
283	22	78.6	487	2	T22825	hypothetical prote	356	22	78.6	5069	2	T17464	rifamycin polyketi
284	22	78.6	488	2	D86250	hypothetical prote	357	21	75.0	46	2	B72808	gp66.1 protein - M
285	22	78.6	492	2	T30011	trans-cinnamate 4-	358	21	75.0	73	2	H69460	conserved hypothet
286	22	78.6	505	2	JC1458	steroid 17alpha-mo	359	21	75.0	77	2	S00970	kcrA1 protein - pl
287	22	78.6	508	1	O4CHC7	4-aminobutyrate tr	360	21	75.0	78	2	T08483	kleA protein - Ent
288	22	78.6	510	2	S68116	trans-cinnamate 4-	361	21	75.0	80	2	T49980	hypothetical prote
289	22	78.6	517	2	T10857	cell division prot	362	21	75.0	82	2	G97763	hypothetical prote
290	22	78.6	517	2	D75494	GMP synthase [impo	363	21	75.0	83	2	S76010	hypothetical prote
291	22	78.6	520	2	C87450	hypothetical prote	364	21	75.0	88	1	EDBE51	immediate-early-5
292	22	78.6	521	2	C87474	phosphate ABC tran	365	21	75.0	95	2	D83360	hypothetical prote
293	22	78.6	533	2	AG2293	metallo-beta-lacta	366	21	75.0	99	2	A69514	hypothetical prote
294	22	78.6	548	2	G82286	conserved hypothet	367	21	75.0	102	1	R3HS15	probable retroelem
295	22	78.6	553	2	D95071	hypothetical prote	368	21	75.0	104	2	T00845	H+-transporting tw
296	22	78.6	553	2	B97939	hypothetical prote	369	21	75.0	106	2	T47202	AcOrf-5 protein -
297	22	78.6	555	2	AB1254	hypothetical prote	370	21	75.0	109	2	E72850	orf5 homolog orf13
298	22	78.6	555	2	AH1616	hypothetical prote	371	21	75.0	109	2	T41895	hypothetical prote
299	22	78.6	557	2	C83130	probable sulfite o	372	21	75.0	110	2	H85234	hypothetical prote
300	22	78.6	573	2	T25397	hypothetical prote	373	21	75.0	110	2	T10608	hypothetical prote
301	22	78.6	575	1	WQEBPI	phosphotransferase	374	21	75.0	112	2	F70403	H+-transporting tw
302	22	78.6	584	2	E83767	ABC transporter (A	375	21	75.0	113	2	A36493	hypothetical prote
303	22	78.6	593	2	S26696	alkaline proteinas	376	21	75.0	115	2	E97483	H+-transporting Ar
304	22	78.6	624	1	BHTLE	hemocyanin chain e	377	21	75.0	120	2	B84364	conserved hypothet
305	22	78.6	637	2	G84356	hypothetical prote	378	21	75.0	129	2	AG0127	arsenate reductase
306	22	78.6	655	2	E82712	peptidyl-prolyl ci	379	21	75.0	130	2	A53641	arsenate reductase
307	22	78.6	660	2	T20569	hypothetical prote	380	21	75.0	131	1	C41902	arsenate reductase
308	22	78.6	662	2	T20570	hypothetical prote	381	21	75.0	131	1	D41903	arsenate reductase
309	22	78.6	687	2	T08528	probable DNA topoi	382	21	75.0	132	2	S43488	homeotic protein L
310	22	78.6	699	2	S67773	hypothetical prote	383	21	75.0	136	2	T22240	hypothetical prote
311	22	78.6	714	2	E81751	transcription elon	384	21	75.0	138	1	QQVLS	gene X protein - g
312	22	78.6	714	2	AF2458	hypothetical prote	385	21	75.0	138	2	H75120	translation initia
313	22	78.6	735	2	S37903	probable ATP-depen	386	21	75.0	138	2	G71010	arsenate reductase
314	22	78.6	755	2	S58718	probable nuclear p	387	21	75.0	138	2	F71010	arsenate reductase
315	22	78.6	768	2	AE3505	alkaline phosphata	388	21	75.0	139	2	F84024	hypothetical prote
316	22	78.6	797	2	S53590	hypothetical prote	389	21	75.0	140	2	A84284	hypothetical prote
317	22	78.6	804	2	H75549	glycogen debranchi	390	21	75.0	141	1	QQVLC1	gene X protein - w
318	22	78.6	826	2	T01494	trehalose-6-phosph	391	21	75.0	141	2	AH2767	conserved hypothet
319	22	78.6	872	2	B86602	alanyl tRNA synthe	392	21	75.0	141	2	C97548	hypothetical prote
320	22	78.6	872	2	H72023	alanine-tRNA ligas	393	21	75.0	144	2	B70360	arsenate reductase
321	22	78.6	875	2	S27572	inner-membrane pro	394	21	75.0	146	2	T08344	hypothetical prote

395	21	75.0	146	2	E90556	hypothetical prote	468	21	75.0	268	2	D85548	probable metal res
396	21	75.0	148	2	D64386	ribosomal protein	469	21	75.0	268	2	B90698	probable metal res
397	21	75.0	149	2	AF2977	conserved hypothet	470	21	75.0	268	2	B64780	probable transport
398	21	75.0	150	2	A83217	conserved hypothet	471	21	75.0	268	2	T23665	hypothetical prote
399	21	75.0	160	2	F98305	hypothetical 16.6K	472	21	75.0	268	2	AG1225	cobalt transport A
400	21	75.0	161	2	S66152	probable DNA-direc	473	21	75.0	268	2	AI1578	cobalt transport A
401	21	75.0	163	2	C85582	unknown protein en	474	21	75.0	269	2	D86739	teichoic acid ABC
402	21	75.0	163	2	H90731	hypothetical prote	475	21	75.0	273	2	E87347	flagellin FljM [im
403	21	75.0	164	2	T30991	hypothetical prote	476	21	75.0	273	2	F87347	flagellin FljN [im
404	21	75.0	165	2	T47011	hypothetical prote	477	21	75.0	273	2	F87430	flagellin FljK [im
405	21	75.0	165	2	AF0237	probable thioredox	478	21	75.0	274	2	G84218	cysteine synthase
406	21	75.0	167	2	E81028	conserved hypothet	479	21	75.0	274	2	S20155	MEF8 protein - yea
407	21	75.0	167	2	A81973	hypothetical prote	480	21	75.0	275	2	C70924	conserved hypothet
408	21	75.0	167	2	AE0300	conserved hypothet	481	21	75.0	276	2	E87682	methylytransferase,
409	21	75.0	168	2	B84211	hypothetical prote	482	21	75.0	277	1	QOECAD	probable membrane
410	21	75.0	172	2	AG1103	B. subtilis yach p	483	21	75.0	277	2	G75518	probable beta-lact
411	21	75.0	178	2	S62846	H+-transporting tw	484	21	75.0	277	2	AC0704	conserved hypothet
412	21	75.0	179	2	H64093	ribosomal protein	485	21	75.0	277	2	F85778	hypothetical prote
413	21	75.0	180	2	A69387	fumarase (fum-1) h	486	21	75.0	277	2	B90930	hypothetical prote
414	21	75.0	180	2	F75607	arsenate reductase	487	21	75.0	279	2	C75491	probable serine es
415	21	75.0	182	2	I64029	hypothetical prote	488	21	75.0	281	2	A84259	hypothetical prote
416	21	75.0	183	1	TVHUR2	transforming prote	489	21	75.0	285	2	G85605	unknown in ISEc8 [
417	21	75.0	183	2	S03180	transforming prote	490	21	75.0	287	2	S25259	streptomycin resis
418	21	75.0	186	2	I46085	proto-oncogene - c	491	21	75.0	288	2	S57184	aryl-alcohol dehyd
419	21	75.0	188	2	T19507	hypothetical prote	492	21	75.0	290	2	A87312	coproporphyrinogen
420	21	75.0	188	2	D89819	conserved hypothet	493	21	75.0	290	2	A84011	transcription acti
421	21	75.0	191	2	S09635	pabA protein - Ser	494	21	75.0	290	2	A69779	conserved hypothet
422	21	75.0	197	2	AF1742	protein involved i	495	21	75.0	292	2	H64149	hypothetical prote
423	21	75.0	197	2	AH1372	proteins involved	496	21	75.0	294	2	AB2367	dTDP-6-deoxy-L-man
424	21	75.0	197	2	AC3447	outer membrane pro	497	21	75.0	294	2	C69956	phosphate ABC tran
425	21	75.0	200	2	T42678	hypothetical prote	498	21	75.0	297	2	A84767	hypothetical prote
426	21	75.0	201	2	AI3639	hypothetical prote	499	21	75.0	297	2	S66102	protein secretion
427	21	75.0	202	2	B83761	hypothetical prote	500	21	75.0	298	2	D95199	primosomal protein
428	21	75.0	204	2	JC4261	somatotropin precu	501	21	75.0	298	2	B98066	primosome componen
429	21	75.0	206	2	AC1823	hypothetical prote	502	21	75.0	304	2	A96642	hypothetical prote
430	21	75.0	207	2	F87470	transcription regu	503	21	75.0	305	2	S19797	FMR1 protein - Pod
431	21	75.0	211	2	T42603	gene 60 protein -	504	21	75.0	307	2	D95363	probable LysR -typ
432	21	75.0	212	1	WZBEF2	gene 60 protein -	505	21	75.0	310	2	H97338	dioxygenase relate
433	21	75.0	212	2	S56199	hypothetical prote	506	21	75.0	314	2	D87576	oxidoreductase, al
434	21	75.0	214	2	E83431	type III export pr	507	21	75.0	315	2	S58171	bifunctional cycla
435	21	75.0	214	2	A47715	cardiac-specific h	508	21	75.0	315	2	T10613	hypothetical prote
436	21	75.0	219	2	AC3026	transcription regu	509	21	75.0	316	2	B90282	hypothetical prote
437	21	75.0	222	2	S77550	hypothetical prote	510	21	75.0	317	2	AI0640	flagellar hook-ass
438	21	75.0	224	2	D75150	hypothetical prote	511	21	75.0	317	2	S10362	hook-associated pr
439	21	75.0	224	2	D70665	probable urease ac	512	21	75.0	318	2	D75210	hypothetical prote
440	21	75.0	226	2	AB1741	ABC-transporter AT	513	21	75.0	324	2	T11971	acetyl-CoA carboxy
441	21	75.0	226	2	AD1371	ABC-transporter AT	514	21	75.0	325	2	A75330	thioredoxin reduct
442	21	75.0	229	2	H75594	hypothetical prote	515	21	75.0	325	2	T48873	electron transfer
443	21	75.0	232	2	C69822	glycine betaine/L-	516	21	75.0	327	2	F85708	unknown protein en
444	21	75.0	233	2	S23522	transcription init	517	21	75.0	329	2	S67807	hypothetical prote
445	21	75.0	234	2	F95127	hypothetical prote	518	21	75.0	330	2	A82538	pseudouridylate sy
446	21	75.0	234	2	C97998	conserved hypothet	519	21	75.0	331	2	AF3267	peptidyl-prolyl ci
447	21	75.0	237	2	AF1566	2-C-methyl-D-eryth	520	21	75.0	333	2	T48363	histidyl-tRNA synt
448	21	75.0	237	2	G84299	hypothetical prote	521	21	75.0	333	2	D83585	hypothetical prote
449	21	75.0	238	2	F98258	hypothetical trans	522	21	75.0	334	2	F82149	Holliday junction
450	21	75.0	238	2	T29797	hypothetical prote	523	21	75.0	334	2	T19860	hypothetical prote
451	21	75.0	242	2	E69450	spermidine/putresc	524	21	75.0	335	2	T07517	DNA-directed RNA p
452	21	75.0	242	2	D90796	hypothetical prote	525	21	75.0	338	2	S67158	26S proteasome reg
453	21	75.0	243	2	T26653	hypothetical prote	526	21	75.0	340	2	AC0459	porphobilinogen sy
454	21	75.0	244	2	G82266	pseudouridine synt	527	21	75.0	341	2	E87390	transcription regu
455	21	75.0	244	2	D86275	F7A19.26 protein -	528	21	75.0	342	2	T08474	kfrA protein - Ent
456	21	75.0	245	2	AC2540	two-component resp	529	21	75.0	343	2	S74937	hypothetical prote
457	21	75.0	246	2	T36203	probable integral	530	21	75.0	346	1	RQPSAA	RecA protein PA361
458	21	75.0	252	2	S74565	ABC-type transport	531	21	75.0	346	2	C81446	homolog of E. coli
459	21	75.0	254	2	T31709	hypothetical prote	532	21	75.0	346	2	H95406	conserved hypothet
460	21	75.0	256	2	A72352	hypothetical prote	533	21	75.0	347	2	E64691	rod shape-determin
461	21	75.0	257	2	E89124	protein K07C11.1 [534	21	75.0	347	2	E71826	rod shape-determin
462	21	75.0	257	2	B75099	hypothetical prote	535	21	75.0	347	2	S22617	hypothetical prote
463	21	75.0	258	2	T34601	hypothetical prote	536	21	75.0	348	2	C87313	transcription regu
464	21	75.0	259	2	AB0565	probable membrane	537	21	75.0	348	2	H90850	hypothetical prote
465	21	75.0	262	2	S63648	H+-transporting tw	538	21	75.0	349	2	E84400	dihydroorotate deh
466	21	75.0	262	2	T18886	hypothetical prote	539	21	75.0	349	2	JN0321	recombination prot
467	21	75.0	264	2	T26558	hypothetical prote	540	21	75.0	349	2	S22448	FMR1 protein - Pod

541	21	75.0	349	2	H85626	hypothetical prote	614	21	75.0	434	2	F75425	trNA nucleotidyltr
542	21	75.0	349	2	B90819	hypothetical prote	615	21	75.0	435	2	T45974	phenylalanine-trNA
543	21	75.0	349	2	E90912	hypothetical prote	616	21	75.0	435	2	A84828	hypothetical prote
544	21	75.0	350	2	B88691	protein F41H10.10	617	21	75.0	436	2	A11316	hypothetical prote
545	21	75.0	355	2	JC1249	peroxidase (EC 1.1	618	21	75.0	436	2	A11688	hypothetical prote
546	21	75.0	359	2	F83850	hypothetical prote	619	21	75.0	438	2	JC1179	transforming prote
547	21	75.0	361	2	G95410	probable ABC trans	620	21	75.0	439	1	TVCTMC	transforming prote
548	21	75.0	363	2	S19419	hypothetical prote	621	21	75.0	439	1	TVHUM	transforming prote
549	21	75.0	364	2	S34355	peroxidase (EC 1.1	622	21	75.0	439	1	TVMS	transforming prote
550	21	75.0	364	2	T43361	probable potassium	623	21	75.0	439	1	TVRTMC	transforming prote
551	21	75.0	365	2	JU0319	hypothetical 43.5K	624	21	75.0	439	2	JC1178	transforming prote
552	21	75.0	368	2	AI1934	hypothetical prote	625	21	75.0	439	2	JU0449	c-myc protein - ch
553	21	75.0	370	2	F70484	twitching mobility	626	21	75.0	440	4	TVHUT	transforming prote
554	21	75.0	370	2	AC1275	aminopeptidase hom	627	21	75.0	441	2	AD0572	probable pyridine
555	21	75.0	370	2	AC1638	aminopeptidase hom	628	21	75.0	442	2	G86294	T24D18.11 protein
556	21	75.0	370	2	A71143	hypothetical prote	629	21	75.0	442	1	QBYPT	hypothetical prote
557	21	75.0	373	1	KMECTD	chorismate mutase	630	21	75.0	443	2	I39538	alpha-amylase - Ae
558	21	75.0	373	1	S29934	chorismate mutase	631	21	75.0	443	2	H96841	hypothetical prote
559	21	75.0	373	2	G91061	chorismate mutase-	632	21	75.0	443	2	C83421	hypothetical prote
560	21	75.0	373	2	C85906	chorismate mutase-	633	21	75.0	444	2	B82296	sigma-54 dependent
561	21	75.0	373	2	AB0399	prephenate dehydro	634	21	75.0	444	2	D97977	glucose inhibited
562	21	75.0	373	2	AI0832	prephenate dehydro	635	21	75.0	445	2	E72674	hypothetical prote
563	21	75.0	373	2	S47455	mitosis protein JN	636	21	75.0	445	2	A83763	L-arabinose transp
564	21	75.0	375	2	E82292	chorismate mutase/	637	21	75.0	447	2	G95068	cysteiny1-trNA syn
565	21	75.0	376	2	S51335	probable aryl-alco	638	21	75.0	447	2	G97936	cysteine-trNA liga
566	21	75.0	376	2	T10594	hypothetical prote	639	21	75.0	448	2	C96542	hypothetical prote
567	21	75.0	377	2	D86881	pyridine nucleotid	640	21	75.0	449	2	H70526	probable cytochrom
568	21	75.0	378	2	E84806	probable elongatio	641	21	75.0	450	2	H64756	probable mercury(I
569	21	75.0	379	1	JC4289	cell division prot	642	21	75.0	450	2	B85522	probable oxidoredu
570	21	75.0	380	1	E69541	conserved hypothet	643	21	75.0	450	2	F90671	probable membrane
571	21	75.0	381	2	T46827	phenoxylbenzoate di	644	21	75.0	450	2	AG0008	probable polyprotei
572	21	75.0	384	2	H89873	hypothetical prote	645	21	75.0	451	1	TVFV2C	aryl hydrocarbon r
573	21	75.0	385	2	E87355	peptide chain rele	646	21	75.0	451	2	T42397	Similar to Flavono
574	21	75.0	385	2	H95413	probable integrase	647	21	75.0	452	2	F96672	two component sens
575	21	75.0	387	2	AC0887	probable alcohol d	648	21	75.0	452	2	AC3136	beta-alanine-pyruv
576	21	75.0	387	2	T09086	sedoheptulose-bisp	649	21	75.0	452	2	AH3525	hypothetical prote
577	21	75.0	387	2	A33497	prophage DLP12 int	650	21	75.0	453	2	T19440	conserved hypothet
578	21	75.0	389	2	AI1506	conserved hypothet	651	21	75.0	456	1	E71284	Gid protein limpor
579	21	75.0	389	2	H69048	conserved hypothet	652	21	75.0	456	2	A95109	tldD-related prote
580	21	75.0	389	2	T16326	hypothetical prote	653	21	75.0	458	1	E71033	argininosuccinate
581	21	75.0	390	2	AI1147	conserved hypothet	654	21	75.0	458	2	B81175	probable argininos
582	21	75.0	396	2	B82320	membrane-bound lyt	655	21	75.0	458	2	C81930	cysteiny1-trNA syn
583	21	75.0	398	2	S65782	serpin - wheat	656	21	75.0	460	2	G83421	hypothetical prote
584	21	75.0	399	2	T06488	serpin WZS2 - whea	657	21	75.0	460	2	T23087	argininosuccinate
585	21	75.0	399	2	A25470	cytokeratin 19 - b	658	21	75.0	461	2	AH2291	cytochrome P450 ho
586	21	75.0	401	2	D87258	fatty oxidation co	659	21	75.0	464	2	T07859	peptidyl-prolyl ci
587	21	75.0	401	2	H87411	acyl-CoA dehydroge	660	21	75.0	464	2	C82756	transforming prote
588	21	75.0	405	2	T47595	RING finger protei	661	21	75.0	465	2	S03325	proline-trNA synth
589	21	75.0	405	2	T41712	hypothetical prote	662	21	75.0	465	2	H84198	cysteine-trNA liga
590	21	75.0	405	2	AB0793	probable MR-MLB-fa	663	21	75.0	466	1	C53402	cysteiny1-trNA syn
591	21	75.0	405	2	G91020	probable racemase	664	21	75.0	466	2	G83663	poly(A) polymerase
592	21	75.0	405	2	T49534	hypothetical prote	665	21	75.0	467	2	D83055	hypothetical prote
593	21	75.0	405	2	F85864	probable racemase	666	21	75.0	468	2	C89912	hypothetical prote
594	21	75.0	405	2	H97130	ftsK-like DNA segr	667	21	75.0	468	2	T10595	hypothetical prote
595	21	75.0	405	2	E64995	hypothetical prote	668	21	75.0	469	2	C71336	hypothetical prote
596	21	75.0	407	2	G95054	ROK family protein	669	21	75.0	471	2	AB3422	glutamy1-trNA(gln)
597	21	75.0	407	2	D97924	xylose repressor p	670	21	75.0	472	2	G70932	probable monoxyme
598	21	75.0	410	2	T36535	probable hydrolase	671	21	75.0	478	2	H71345	probable polynucle
599	21	75.0	411	2	T51818	ATP phosphoribosyl	672	21	75.0	481	2	B96691	probable cytochrom
600	21	75.0	413	2	T47272	transposase tnpA [673	21	75.0	484	1	TVMVFT	gag-myc polypotei
601	21	75.0	416	1	FOCH	transforming prote	674	21	75.0	484	2	AF3529	flagellar hook-ass
602	21	75.0	416	1	TVFVAC	transforming prote	675	21	75.0	486	2	T49251	histidine-trNA lig
603	21	75.0	416	2	C96725	hypothetical prote	676	21	75.0	495	2	C69084	phosphoserine phos
604	21	75.0	423	1	TVFV2E	transforming prote	677	21	75.0	501	1	JN0539	head protein gp5 -
605	21	75.0	423	2	G95880	probable trehalose	678	21	75.0	501	2	C85743	probable head-tail
606	21	75.0	425	2	A98152	sensor histidine k	679	21	75.0	501	2	D90767	probable head-tail
607	21	75.0	428	2	AC2288	cell division prot	680	21	75.0	501	2	D90970	hypothetical prote
608	21	75.0	428	4	S46932	hypothetical cell-	681	21	75.0	501	2	D85717	spliceosome-associ
609	21	75.0	430	2	AI0611	seryl-trNA synthet	682	21	75.0	502	2	A55749	probable carboxyle
610	21	75.0	430	2	AE0168	serine-trNA ligase	683	21	75.0	502	2	T35910	lipopolysaccharide
611	21	75.0	430	2	D83977	hypothetical prote	684	21	75.0	502	2	A83938	pollen-specific pr
612	21	75.0	431	2	F89817	conserved hypothet	685	21	75.0	504	2	T07129	unknown protein, 7
613	21	75.0	431	2	H71172	hypothetical prote	686	21	75.0	508	2	F86458	

687 21 75.0 508 2 T22626 hypothetical prote
688 21 75.0 509 2 T00793 hypothetical prote
689 21 75.0 525 2 S69991 alcohol O-acetyltr
690 21 75.0 525 2 S67289 alcohol O-acetyltr
691 21 75.0 526 2 S54025 Grpase MSS1, mitoc
692 21 75.0 528 2 G97736 hypothetical prote
693 21 75.0 529 2 T10612 hypothetical prote
694 21 75.0 535 2 C95057 Crp synthase [limp
695 21 75.0 535 2 F97926 Crp synthase (EC 6
696 21 75.0 538 2 T08640 hypothetical prote
697 21 75.0 541 2 C96492 probable pectinest
698 21 75.0 541 2 T49108 pectinesterase lik
699 21 75.0 542 2 S76358 GMP synthase (glut
700 21 75.0 547 2 G71307 probable oligopept
701 21 75.0 548 2 T05670 pollen-specific pr
702 21 75.0 548 2 A71544 probable diphospha
703 21 75.0 549 2 T02790 hypothetical prote
704 21 75.0 551 1 QQVZ25 rifampicin resista
705 21 75.0 551 2 H86351 protein T26F17.6 [i
706 21 75.0 551 2 A36848 N3L protein - vari
707 21 75.0 551 2 T28541 hypothetical prote
708 21 75.0 551 2 E72163 O3L protein - vari
709 21 75.0 552 2 T33760 sphingosine-1-phos
710 21 75.0 554 2 S22495 pollen-specific pr
711 21 75.0 560 2 S54093 hypothetical prote
712 21 75.0 561 2 T05545 pollen-specific pr
713 21 75.0 565 2 T22649 hypothetical prote
714 21 75.0 569 1 S35291 urease (EC 3.5.1.5
715 21 75.0 570 2 S52765 secD protein - Str
716 21 75.0 575 1 WQECPI phosphotransferase
717 21 75.0 575 1 SYBYMM methionine-tRNA li
718 21 75.0 575 2 C85884 PEP-protein phosph
719 21 75.0 575 2 H91039 PEP-protein phosph
720 21 75.0 577 2 A57469 CMP-N-acetylneuram
721 21 75.0 580 2 T21493 hypothetical prote
722 21 75.0 583 2 S01496 lamin B - African
723 21 75.0 583 2 F70592 probable lpqB prot
724 21 75.0 584 2 S05518 lamin B-1 - chicke
725 21 75.0 584 2 T23174 hypothetical prote
726 21 75.0 586 1 VEHULB lamin B1 - human
727 21 75.0 587 2 T07634 pollen-specific pr
728 21 75.0 587 2 AD3519 2,3-dihydroxybenzo
729 21 75.0 587 2 S07720 lamin B - mouse
730 21 75.0 597 2 D70100 phosphoglucomutase
731 21 75.0 598 1 A37192 excinuclease ABC,
732 21 75.0 600 2 T36195 probable acyl-CoA
733 21 75.0 601 2 T49752 hypothetical prote
734 21 75.0 603 2 T38920 methylenetetrahydr
735 21 75.0 606 2 F69493 conserved hypothet
736 21 75.0 609 1 A42537 gene 16 protein -
737 21 75.0 616 2 G86890 proline-tRNA ligas
738 21 75.0 619 2 S25334 H+-exporting ATPas
739 21 75.0 622 2 JC5425 transcription init
740 21 75.0 629 2 E82624 thiamin biosynthes
741 21 75.0 629 2 T06675 hypothetical prote
742 21 75.0 630 2 T38023 probable transcrip
743 21 75.0 631 2 A53623 yolk protein facto
744 21 75.0 636 2 F69027 cleavage and polya
745 21 75.0 641 2 T49477 phenol hydroxylase
746 21 75.0 643 2 T19199 hypothetical prote
747 21 75.0 644 2 AD3471 adenylyl-sulfate k
748 21 75.0 645 2 S50374 Grp-binding protei
749 21 75.0 651 2 E86242 hypothetical prote
750 21 75.0 652 2 G82401 methyl-accepting c
751 21 75.0 653 2 H97777 DNA helicase II (E
752 21 75.0 653 2 S67035 probable membrane
753 21 75.0 654 2 I56134 tumor necrosis fac
754 21 75.0 658 2 F71703 DNA helicase II (u
755 21 75.0 658 2 T05655 hypothetical prote
756 21 75.0 659 1 A64228 DNA ligase (NAD) (
757 21 75.0 661 2 C83843 hypothetical prote
758 21 75.0 661 2 T15802 hypothetical prote
759 21 75.0 663 2 B70460 excinuclease ABC c

760 21 75.0 668 2 AG1989 hypothetical prote
761 21 75.0 670 2 S53414 probable membrane
762 21 75.0 673 1 BVECUB excinuclease ABC c
763 21 75.0 673 2 AD0597 excision nuclease
764 21 75.0 673 2 B85586 DNA repair, excisi
765 21 75.0 673 2 A99736 excision nuclease
766 21 75.0 678 2 H88187 protein C18H9.8 [i
767 21 75.0 679 2 D89625 excinuclease ABC c
768 21 75.0 684 2 H96646 hypothetical prote
769 21 75.0 684 2 T02149 hypothetical prote
770 21 75.0 692 2 A70933 probable nrdZ prot
771 21 75.0 693 2 C70167 translation elonga
772 21 75.0 699 2 T16109 hypothetical prote
773 21 75.0 699 2 T18426 hypothetical prote
774 21 75.0 701 2 JC8062 L-glutamate oxidas
775 21 75.0 701 2 S35313 TIP1 protein - yea
776 21 75.0 701 2 T28151 probable ABC-type
777 21 75.0 709 2 T00096 DNA replication-re
778 21 75.0 712 2 T33028 hypothetical prote
779 21 75.0 719 2 AI0769 protein-tyrosine k
780 21 75.0 724 2 AE1199 ATP-dependent prot
781 21 75.0 724 2 AC1557 probable exported
782 21 75.0 728 2 AE0101 excinuclease ABC c
783 21 75.0 730 2 G75292 GGDEF family prote
784 21 75.0 734 2 AF3108 hypothetical prote
785 21 75.0 737 2 C70770 probable chemotaxi
786 21 75.0 739 2 F71161 hypothetical prote
787 21 75.0 742 2 E89801 hypothetical prote
788 21 75.0 746 2 T19287 nitrogen fixation
789 21 75.0 747 2 F98178 probable competenc
790 21 75.0 754 2 AE0614 hypothetical prote
791 21 75.0 758 2 F72363 probable thyroid r
792 21 75.0 761 2 T03719 ribonucleoside-dip
793 21 75.0 775 1 WMBE19 F13B4.4 protein -
794 21 75.0 786 2 E86268 aculeacin-A acylas
795 21 75.0 786 2 JC1298 formate C-acetyltr
796 21 75.0 787 2 H86707 probable ubiquinol
797 21 75.0 788 2 S48191 leucyl-trna synthe
798 21 75.0 807 2 E90523 probable transmemb
799 21 75.0 811 2 T36581 kinesin-like prote
800 21 75.0 813 2 T46242 hypothetical prote
801 21 75.0 817 2 T04697 DNA mismatch repai
802 21 75.0 822 2 S77112 hypothetical prote
803 21 75.0 830 2 T16236 BR01 protein - yea
804 21 75.0 844 2 S61104 aryl hydrocarbon r
805 21 75.0 846 2 JC7721 probable trehalose
806 21 75.0 862 2 E84567 PAL1 protein - yea
807 21 75.0 870 2 S65158 preprotein translo
808 21 75.0 871 2 C72238 DNA-directed DNA p
809 21 75.0 872 2 JC7380 hypothetical prote
810 21 75.0 873 2 E90581 gag-myc polyprotei
811 21 75.0 875 1 FOFV29 gag-myc polyprotei
812 21 75.0 877 2 T35861 probable large sec
813 21 75.0 884 2 E71289 probable penicilli
814 21 75.0 884 2 B96598 hypothetical prote
815 21 75.0 887 2 B96598 translation initia
816 21 75.0 890 2 E81576 translation factor-
817 21 75.0 890 2 E86530 translation initia
818 21 75.0 890 2 F72093 translation initia
819 21 75.0 896 2 G81709 isoleucyl-tRNA syn
820 21 75.0 900 2 G82894 hypothetical prote
821 21 75.0 903 2 T47316 two-component sens
822 21 75.0 911 2 AD0959 hypothetical prote
823 21 75.0 925 2 T29311 glutamate-ammonia-
824 21 75.0 948 2 C82077 excinuclease ABC (
825 21 75.0 957 2 B84099 DNA translocase (s
826 21 75.0 960 2 B84056 DNA-directed RNA p
827 21 75.0 976 2 D96714 hypothetical prote
828 21 75.0 982 2 T04542 bacteriophage N4 a
829 21 75.0 990 2 B49351 bacteriophage N4 a
830 21 75.0 990 2 H90703 bacteriophage N4 a
831 21 75.0 990 2 C85554 bacteriophage N4 a
832 21 75.0 990 2 C85554 bacteriophage N4 a

833 21 75.0 991 2 A99334 hypothetical prote
834 21 75.0 1011 2 T51399 DNA-directed RNA p
835 21 75.0 1013 2 T30818 hsdR protein - Kle
836 21 75.0 1019 2 A38738 coagulation factor
837 21 75.0 1021 2 T05901 hypothetical prote
838 21 75.0 1023 2 AG3023 AcrB/AcrD/AcrF fam
839 21 75.0 1023 2 B98261 probable rnd efflu
840 21 75.0 1031 2 B81288 probable sugar tra
841 21 75.0 1045 2 B30239 hydroxymethylgluta
842 21 75.0 1085 2 JC2227 probable helicase
843 21 75.0 1088 1 P1XRPR inner layer protei
844 21 75.0 1088 2 S39261 VP1 protein - porc
845 21 75.0 1100 2 G83376 probable trehalose
846 21 75.0 1143 2 S46122 SNF2 protein homol
847 21 75.0 1145 2 T05573 hypothetical prote
848 21 75.0 1148 2 F69685 pyruvate carboxyla
849 21 75.0 1170 2 A72287 hypothetical prote
850 21 75.0 1177 2 AD0438 probable exported
851 21 75.0 1204 2 B81947 probable exodeoxyr
852 21 75.0 1207 2 C70013 conserved hypothet
853 21 75.0 1226 2 T49915 pre-mRNA splicing
854 21 75.0 1235 2 S24109 phosphorylase kina
855 21 75.0 1236 2 E70977 hypothetical prote
856 21 75.0 1277 2 S70306 hypothetical prote
857 21 75.0 1309 1 BVBYD9 RAD9 protein - yea
858 21 75.0 1321 2 T42228 P-glycoprotein sis
859 21 75.0 1321 2 T42842 bile salt transpor
860 21 75.0 1383 2 T06091 hypothetical prote
861 21 75.0 1394 2 S66876 ATP-dependent tran
862 21 75.0 1396 2 G71529 DNA-directed RNA p
863 21 75.0 1396 2 T10627 hypothetical prote
864 21 75.0 1449 2 B84426 retrotransposon li
865 21 75.0 1474 2 B85188 hypothetical prote
866 21 75.0 1479 2 A89793 B. subtilis Yuka p
867 21 75.0 1498 2 AF1082 B. subtilis Yuka p
868 21 75.0 1498 2 AG1439 hypothetical prote
869 21 75.0 1509 2 T19486 hypothetical prote
870 21 75.0 1525 2 T14961 F9H16.4 protein -
871 21 75.0 1551 2 F86342 hypothetical prote
872 21 75.0 1580 2 T26204 hypothetical prote
873 21 75.0 1581 2 B71636 hypothetical prote
874 21 75.0 1622 2 T45240 probable membrane
875 21 75.0 1648 2 S61654 probable purine nu
876 21 75.0 1758 2 S57015 hypothetical prote
877 21 75.0 1780 2 T17272 hypothetical prote
878 21 75.0 1864 2 T18485 probable 1-phospha
879 21 75.0 1900 2 S45530 hypothetical prote
880 21 75.0 2088 2 E71436 hypothetical prote
881 21 75.0 2265 2 T26183 hypothetical prote
882 21 75.0 2697 2 T25444 hypothetical prote
883 21 75.0 2748 2 S57976 nuclear migration
884 21 75.0 3206 1 GNVSPV genome polypeptin
885 21 75.0 3227 2 T37964 probable ubiquitin
886 21 75.0 3228 2 T21381 hypothetical prote
887 21 75.0 3434 1 GNWVMV genome polypeptin
888 21 75.0 3491 2 T43231 probable 6-deoxyer
889 21 75.0 3591 1 S21010 filamentous hemagg
890 21 75.0 3635 2 T10053 laminin alpha 5 ch
891 21 75.0 4385 2 T29042 hypothetical prote
892 21 75.0 4967 2 S72269 ryanodine receptor
893 21 75.0 4969 2 A37113 ryanodine receptor
894 21 75.0 7576 2 T17428 PK506 polyketide s
895 20 71.4 15 2 S57577 T cell receptor V-
896 20 71.4 27 2 A61412 methane monooxygen
897 20 71.4 60 2 F87301 hypothetical prote
898 20 71.4 62 2 S57781 hypothetical prote
899 20 71.4 62 2 AI2130 hypothetical prote
900 20 71.4 68 2 T10328 hypothetical prote
901 20 71.4 75 2 E70870 hypothetical prote
902 20 71.4 76 2 F83566 hypothetical prote
903 20 71.4 78 2 T35577 hypothetical prote
904 20 71.4 83 2 S28120 gas-vesicle operon
905 20 71.4 84 2 B69014 ferredoxin 2[4Fe-4

906 20 71.4 85 2 D87324 hypothetical prote
907 20 71.4 87 2 B69932 hypothetical prote
908 20 71.4 90 2 T42122 probable transposa
909 20 71.4 93 2 A95121 hypothetical prote
910 20 71.4 93 2 F97990 hypothetical prote
911 20 71.4 95 2 T45336 hypothetical prote
912 20 71.4 95 2 AB3126 hypothetical prote
913 20 71.4 96 1 BUKV basic blue protein
914 20 71.4 96 2 AH0115 conserved hypothet
915 20 71.4 96 2 D83228 hypothetical prote
916 20 71.4 97 2 E86682 prophage pi1 prote
917 20 71.4 97 2 A86754 prophage pi2 prote
918 20 71.4 98 2 AD1146 methyltransferase
919 20 71.4 98 2 AD1505 methyltransferase
920 20 71.4 101 2 T36356 probable ribosomal
921 20 71.4 101 2 C83539 morphogene protein
922 20 71.4 101 2 S75079 hypothetical prote
923 20 71.4 102 2 H86160 hypothetical prote
924 20 71.4 102 2 B84355 30S ribosomal prot
925 20 71.4 102 2 A69497 periplasmic divale
926 20 71.4 103 2 D97247 hypothetical prote
927 20 71.4 104 2 S12188 hypothetical prote
928 20 71.4 105 2 AI0557 BOLA protein limpo
929 20 71.4 106 2 AH0383 BOLA protein limpo
930 20 71.4 106 2 G82795 hypothetical prote
931 20 71.4 110 2 S29724 lignin peroxidase
932 20 71.4 113 2 D70580 hypothetical prote
933 20 71.4 114 2 JC5238 galactosylceramide
934 20 71.4 115 2 AB0657 integrase (truncat
935 20 71.4 115 2 B85883 probable transcrip
936 20 71.4 116 2 H45893 T-cell receptor al
937 20 71.4 116 2 S17567 AQN-3 protein - pi
938 20 71.4 116 2 S39434 spermadhesin AQN-3
939 20 71.4 116 2 T51015 hypothetical prote
940 20 71.4 116 2 H91038 hypothetical prote
941 20 71.4 117 2 I68524 ribosomal protein
942 20 71.4 118 2 G64302 hypothetical prote
943 20 71.4 119 2 S53460 hypothetical prote
944 20 71.4 120 2 E81229 30S ribosomal prot
945 20 71.4 120 2 D69833 hypothetical prote
946 20 71.4 121 2 A41940 Ig heavy chain V r
947 20 71.4 123 2 G64219 ribosomal protein
948 20 71.4 124 2 S62816 phosphate acceptor
949 20 71.4 124 2 S20545 hypothetical prote
950 20 71.4 124 2 S11769 hypothetical prote
951 20 71.4 126 2 S35338 MLC536.21c protei
952 20 71.4 126 2 T11008 transposase all269
953 20 71.4 127 2 AF2142 transposase all440
954 20 71.4 127 2 AH2335 transposase all481
955 20 71.4 127 2 AI2407 transposase all443
956 20 71.4 127 2 AF2360 transposase alr515
957 20 71.4 127 2 AE2450 transposase all001
958 20 71.4 127 2 AH1808 transposase alr361
959 20 71.4 127 2 AC2257 hypothetical prote
960 20 71.4 128 2 E70547 lactoylglutathione
961 20 71.4 128 2 B97170 antigen B (importe
962 20 71.4 129 2 AF1089 arsenate reductase
963 20 71.4 130 2 B69047 hypothetical prote
964 20 71.4 131 2 T01293 T-cell receptor al
965 20 71.4 132 2 B31211 hypothetical prote
966 20 71.4 132 2 T02116 hypothetical prote
967 20 71.4 132 2 B83299 hypothetical prote
968 20 71.4 133 2 AD1875 probable small pro
969 20 71.4 134 2 B71354 conserved hypothet
970 20 71.4 134 2 F87396 gene lmdmri protei
971 20 71.4 135 2 B44438 sperm motility inh
972 20 71.4 136 2 S72508 hypothetical prote
973 20 71.4 136 2 G70540 probable heavy met
974 20 71.4 137 2 T36635 hypothetical prote
975 20 71.4 137 2 T17718 hypothetical prote
976 20 71.4 137 2 AI3194 hypothetical prote
977 20 71.4 138 2 AB0538 probable secreted

979	20	71.4	138	2	C87389	hypothetical prote
980	20	71.4	138	2	A87192	conserved hypothet
981	20	71.4	140	2	B87622	cytidine deaminase
982	20	71.4	141	2	S14259	seed storage prote
983	20	71.4	142	1	E64794	ybdQ protein - Esc
984	20	71.4	142	2	E70356	hypothetical prote
985	20	71.4	142	2	F90709	hypothetical prote
986	20	71.4	142	2	B85560	hypothetical prote
987	20	71.4	142	2	H90468	hypothetical prote
988	20	71.4	146	2	AD0524	probable PTS syste
989	20	71.4	147	2	C64223	hypothetical prote
990	20	71.4	147	2	G70209	conserved hypothet
991	20	71.4	148	2	A26192	pol polyprotein -
992	20	71.4	148	2	C71829	hypothetical prote
993	20	71.4	148	2	B64687	conserved hypothet
994	20	71.4	150	2	B71329	probable flagellar
995	20	71.4	151	2	H88480	protein Cl6A3.8 [i
996	20	71.4	151	2	A69195	transcription regu
997	20	71.4	151	2	T19087	hypothetical prote
998	20	71.4	152	2	D72341	purine-binding che
999	20	71.4	152	2	AF2680	transcription regu
1000	20	71.4	152	2	D97462	hypothetical prote

ALIGNMENTS

RESULT 1
T02022
hypothetical protein T9E19.1 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C;Accession: T02022
R;Strowmatt, C.; Johnson, D.; Le, T.
submitted to the EMBL Data Library, November 1998
A;Description: The sequence of A. thaliana T9E19.
A;Reference number: Z14496
A;Accession: T02022
A;Status: translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-543 <STR>
A;Cross-references: UNIPROT:Q9ZSH5; UNIPARC:UPI000009C860; EMBL:AF104920; NID:g3859610;
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 198/3; 253/3; 295/3; 330/3; 364/3; 384/3
A;Note: T9E19.1

Query Match	96.4%;	Score 27;	DB 2;	Length 543;
Best Local Similarity	66.7%;	Pred. No. 67;		
Matches	6;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 69 VRDQIRAAAL 77

RESULT 2
T47381
hypothetical protein T5C2.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47381
R;Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24463
A;Accession: T47381
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1113 <OBE>
A;Cross-references: UNIPROT:Q9M180; UNIPARC:UPI00000A5EDF; EMBL:AL138664
A;Experimental source: cultivar Columbia; BAC clone T5C2
C;Genetics:

A;Map position: 3
A;Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 732/3; 767/3; 801/3; 816/3; 845/3; 9
A;Note: T5C2.90

Query Match	96.4%;	Score 27;	DB 2;	Length 1113;
Best Local Similarity	66.7%;	Pred. No. 1.5e+02;		
Matches	6;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 488 VRDQIRAAAL 496

RESULT 3
E86402
hypothetical protein F28L5.2 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86402
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86402
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1198 <STO>
A;Cross-references: UNIPROT:Q9C6N9; UNIPARC:UPI00000A5394; GB:AE005172; NID:g10998930; P-
C;Genetics:
A;Map position: 1

Query Match	96.4%;	Score 27;	DB 2;	Length 1198;
Best Local Similarity	66.7%;	Pred. No. 1.6e+02;		
Matches	6;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 486 VRDQIRATL 494

RESULT 4
F86386
hypothetical protein F14G11.4 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86386
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1201 <STO>
A;Cross-references: UNIPROT:Q9C607; UNIPARC:UPI00000A1C16; GB:AE005172; NID:g11560181; P-
C;Genetics:
A;Map position: 1

Query Match	96.4%;	Score 27;	DB 2;	Length 1201;
-------------	--------	-----------	-------	--------------

Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 455 VRDQIRAAAL 463

RESULT 5
H96559
hypothetical protein F5F19.8 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H96559
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1305 <STO>
A;Cross-references: UNIPROT:Q9ZU21; UNIPARC:UPI00000AA103; GB:AE005173; NID:g4220449; PI
C;Genetics:
A;Gene: F5F19.8
A;Map position: 1

Query Match 96.4%; Score 27; DB 2; Length 1305;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 491 VRDQIRAAAL 499

RESULT 6
A84470
hypothetical protein At2g05560 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84470
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84470
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1472 <STO>
A;Cross-references: UNIPROT:Q9SL11; UNIPARC:UPI00000A7470; GB:AE002093; NID:g4581168; PI
C;Genetics:
A;Gene: At2g05560
A;Map position: 2

Query Match 96.4%; Score 27; DB 2; Length 1472;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 403 VRDQIRAAAL 411

RESULT 7
T02599
hypothetical protein At2g14770 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F26C24.9
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02599; C84521
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A;Reference number: Z14680
A;Accession: T02599
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1756 <ROU>
A;Cross-references: UNIPROT:O80975; UNIPARC:UPI00000AA9BA; EMBL:AC004705; NID:g3252804;
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1756 <STO>
A;Cross-references: UNIPARC:UPI00000AA9BA; GB:AE002093; NID:g3252818; PIDN:AAC24188.1; C
C;Genetics:
A;Gene: At2g14770; F26C24.9
A;Map position: 2
A;Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 725/1; 758/3; 778/3; 812/3; 902/2; 9

Query Match 96.4%; Score 27; DB 2; Length 1756;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 488 VRDQIRAAAL 496

RESULT 8
T01194
hypothetical protein F21E10.9 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01194
R;Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F21E10.
A;Reference number: Z14258
A;Accession: T01194
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <DAV>
A;Cross-references: UNIPROT:O65247; UNIPARC:UPI00000A324E; EMBL:AF058914; NID:g3047074;
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: ATSP:F21E10.9
A;Map position: 5

Query Match 92.9%; Score 26; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 187 VRDQIQITL 195

RESULT 9
AD1062

protein kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD1062
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD1062
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <PAR>
A;Cross-references: UNIPARC:UPI000005A993; GB:AL513382; PIDN:CAD06945.1; PID:g16505592;
C;Genetics:
A;Gene: STY4823

Query Match 92.9%; Score 26; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 255 VRDQIAKYL 263

RESULT 10
B36329
hypothetical protein 2 - cabbage looper transposon TED (fragment)
C;Species: Trichoplusia ni (cabbage looper)
C;Date: 01-Feb-1991 #sequence_revision 01-Feb-1991 #text_change 30-Sep-1993
C;Accession: B36329
R;Friesen, P.D.; Nissen, M.S.
Mol. Cell. Biol. 10, 3067-3077, 1990
A;Title: Gene organization and transcription of TED, a lepidopteran retrotransposon inte
A;Reference number: A36329; MUID:90258898; PMID:1692964
A;Accession: B36329
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1236 <PRI>
A;Cross-references: UNIPARC:UPI000017BE4E; GB:M32662

Query Match 92.9%; Score 26; DB 2; Length 1236;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 306 VRDQITKML 314

RESULT 11
D85066
hypothetical protein AT4g05280 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85066
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: D85066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1312 <STO>
A;Cross-references: UNIPROT:Q9M0W7; UNIPARC:UPI000000A5C54; GB:NC_001268; NID:g7267288; F
C;Genetics:
A;Gene: AT4g05280
A;Map position: 4

Query Match 92.9%; Score 26; DB 2; Length 1312;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 434 VRDQIRAVL 442

RESULT 12
S09866
hypothetical protein ULL01 - human cytomegalovirus (strain AD169)
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S09866
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; K
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09866
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-115 <CHE>
A;Cross-references: UNIPROT:P16826; UNIPARC:UPI0000137C06; EMBL:X17403; NID:G59591; PIDN
A;Note: this sequence was submitted to the EMBL Data Library, December 1989
A;Note: this reading frame extends between two stop codons and does not begin with a star
C;Superfamily: human cytomegalovirus hypothetical protein ULL01

Query Match 89.3%; Score 25; DB 2; Length 115;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 73 VRDQLSSQL 81

RESULT 13
D83620
hypothetical protein PA0201 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: D83620
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-193 <STO>
A;Cross-references: UNIPROT:Q9I6T7; UNIPARC:UPI000000C4F91; GB:AE004458; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0201

Query Match 89.3%; Score 25; DB 2; Length 193;
Best Local Similarity 55.6%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
Db 87 VRDQVRGAL 95

RESULT 14
T23381
hypothetical protein K07A1.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23381
R;Percy, C.
A;Title: Submitted to the EMBL Data Library, October 1996
A;Reference number: Z19734
A;Accession: T23381
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-243 <WIL>
A;Cross-references: UNIPROT:P90912; UNIPARC:UPI0000082F5A; EMBL:Z81097; PIDN:CAB03168.1;
A;Experimental source: clone K07A1
C;Genetics:
A;Gene: CESP:K07A1.7
A;Map position: 1
A;Introns: 27/3; 113/3; 187/1

Query Match 89.3%; Score 25; DB 2; Length 243;
Best Local Similarity 55.6%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||||: |
Db 206 VRDQLATQL 214

RESULT 15
S14268
peroxidase (EC 1.11.1.7), neutral - horseradish
C;Species: Armoracia rusticana (horseradish)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S14268
R;Bartonek-Roxa, E.; Eriksson, H.; Mattiasson, B.
Biochim. Biophys. Acta 1088, 245-250, 1991
A;Title: The cDNA sequence of a neutral horseradish peroxidase.
A;Reference number: S14268; MUID:91159476; PMID:2001399
A;Accession: S14268
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-327 <BAR>
A;Cross-references: UNIPROT:Q42517; UNIPARC:UPI0000131635; GB:X57564; NID:G16095; PIDN:C
C;Superfamily: peroxidase
C;Keywords: glycoprotein; heme; iron; metalloprotein; oxidoreductase
F;39-116/Disulfide bonds: #status predicted
F;66/Active site: Arg #status predicted
F;70,193/Binding site: heme iron (His) (axial ligands) #status predicted
F;72-77/Disulfide bonds: #status predicted
F;122-323/Disulfide bonds: #status predicted
F;200-232/Disulfide bonds: #status predicted

Query Match 89.3%; Score 25; DB 2; Length 327;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||||: |
Db 46 VRDQVKIAL 54

RESULT 16
A25228
protein kinase CDC7 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: cell division control protein CDC7; protein D2855; protein YDL017w
C;Species: Saccharomyces cerevisiae
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 31-Dec-2004
C;Accession: A25228; JT0485; S29519; S52502; S67549
R;Paterson, M.; Sclafani, R.A.; Fangman, W.L.; Rosamond, J.
Mol. Cell. Biol. 6, 1590-1598, 1986
A;Title: Molecular characterization of cell cycle gene CDC7 from Saccharomyces cerevisiae
A;Reference number: A25228; MUID:87064440; PMID:3537706
A;Accession: A25228
A;Molecule type: DNA
A;Residues: 1-507 <PAT>
A;Cross-references: UNIPROT:P06243; UNIPARC:UPI000012725A; EMBL:M12624; NID:G171201; PID

R;Bahman, M.; Buck, V.; White, A.; Rosamond, J.
Biochim. Biophys. Acta 951, 335-343, 1988
A;Title: Characterisation of the CDC7 gene product of Saccharomyces cerevisiae as a prot
A;Reference number: JT0485; MUID:89088247; PMID:2850010
A;Accession: JT0485
A;Molecule type: DNA
A;Residues: 1-138 <BAH>
A;Cross-references: UNIPARC:UPI0000168B5F; EMBL:X14164; NID:G3511; PIDN:CAA32369.1; PID:
R;Ham, J.; Moore, D.; Rosamond, J.; Johnston, I.R.
Nucleic Acids Res. 17, 5781-5792, 1989
A;Title: Transcriptional analysis of the CDC7 protein kinase gene of Saccharomyces cerev
A;Reference number: S29519; MUID:89345168; PMID:2668893
A;Accession: S29519
A;Molecule type: DNA
A;Residues: 1-20 <HAM>
A;Cross-references: UNIPARC:UPI000017A44F; EMBL:X15362
R;Andre, B.; Visser, S.; Urrestarazu, L.
submitted to the EMBL Data Library, February 1995
A;Description: The sequence of a 42 kb segment located on the left arm of chromosome IV
A;Reference number: S52492
A;Accession: S52502
A;Molecule type: DNA
A;Residues: 1-507 <AND>
A;Cross-references: UNIPARC:UPI000012725A; EMBL:Z48432; NID:G683669; PIDN:CAA88342.1; PI
R;Urrestarazu, L.A.; Andre, B.; Visser, S.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67535
A;Accession: S67549
A;Molecule type: DNA
A;Residues: 1-507 <URR>
A;Cross-references: UNIPARC:UPI000012725A; EMBL:Z74065; NID:G1430982; PIDN:CAA98574.1; P
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:CDC7
A;Cross-references: SGD:S0002175; MIPS:YDL017w
A;Map position: 4L
C;Keywords: ATP; cell cycle control; phosphotransferase; serine/threonine-specific prote
F;31-394/Domain: protein kinase homology <KIN>
F;39-47/Region: protein kinase ATP-binding motif

Query Match 89.3%; Score 25; DB 2; Length 507;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
||||: |
Db 112 VRDQVIAVL 120

RESULT 17
F82192
ABC transporter, ATP-binding protein VC1499 [imported] - Vibrio cholerae (strain N16961
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: F82192
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: F82192
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-610 <HEI>
A;Cross-references: UNIPROT:Q9KRY4; UNIPARC:UPI00000C2FE1; GB:AE004228; GB:AE003852; NI
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1499
A;Map position: 1

Query Match 89.3%; Score 25; DB 2; Length 610;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 VRDQIXXXL 9 :
Db	444 IRDQIAVVL 452
RESULT 18	
AF3016 Na+/H+ antiporter Atu3738 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)	
C;Species: Agrobacterium tumefaciens	
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004	
C;Accession: AF3016	
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.	
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaivin, T.; Levy, R.; Li, M.; McClell	
; Karp, P.; Romero, P.; Zhang, S.	
Science 294, 2317-2323, 2001	
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,	
ster, E.W.	
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.	
A;Reference number: AB2577; MUID:21608550; PMID:11743193	
A;Accession: AF3016	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-621 <KUR>	
A;Cross-references: UNIPROT:Q8U9J0; UNIPARC:UPI0000164865; GB:AE008689; PIDN:AAL44548.1;	
A;Experimental source: strain C58 (Dupont)	
C;Genetics:	
A;Gene: Atu3738	
A;Map position: linear chromosome	
Query Match 89.3%; Score 25; DB 2; Length 621;	
Best Local Similarity 44.4%; Pred. No. 2.6e+02;	
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY	1 VRDQIXXXL 9 :
Db	145 IRDQVRSTL 153
RESULT 19	
B98268 probable sodium/hydrogen antiporter PA5021 [imported] - Agrobacterium tumefaciens (strain	
C;Species: Agrobacterium tumefaciens	
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004	
C;Accession: B98268	
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,	
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;	
Science 294, 2323-2328, 2001	
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum	
A;Reference number: A97359; MUID:21608551; PMID:11743194	
A;Accession: B98268	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-642 <KUR>	
A;Cross-references: UNIPROT:Q8U9J0; UNIPARC:UPI00000D22F3; GB:AE007870; PIDN:AAK89668.1;	
C;Genetics:	
A;Gene: AGR_L_2194	
A;Map position: linear chromosome	
Query Match 89.3%; Score 25; DB 2; Length 642;	
Best Local Similarity 44.4%; Pred. No. 2.7e+02;	
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY	1 VRDQIXXXL 9 :
Db	166 IRDQVRSTL 174
RESULT 20	
F89895 ATP-dependent DNA helicase [imported] - Staphylococcus aureus (strain N315)	
C;Species: Staphylococcus aureus	

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004	
C;Accession: F89895	
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi	
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;	
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.	
Lancet 357, 1225-1240, 2001	
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.	
A;Reference number: A89758; MUID:21311952; PMID:11418146	
A;Accession: F89895	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-686 <KUR>	
A;Cross-references: UNIPROT:Q99UP1; UNIPARC:UPI00000CA9C1; GB:BA000018; PID:gl3701027; PFI	
A;Experimental source: strain N315	
C;Genetics:	
A;Gene: recG	
C;Superfamily: DNA helicase recG	
Query Match 89.3%; Score 25; DB 2; Length 686;	
Best Local Similarity 55.6%; Pred. No. 2.9e+02;	
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 VRDQIXXXL 9 :
Db	166 IRDQIRQAL 174
RESULT 21	
T00122 hypothetical 88.9K protein - Leptospira interrogans	
C;Species: Leptospira interrogans	
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004	
C;Accession: T00122	
R;Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.	
Gene 215, 37-45, 1998	
A;Title: Physical and genetic maps of the Leptospira interrogans serovar icterohaemorrhag	
A;Reference number: Z14115; MUID:98332717; PMID:9666070	
A;Accession: T00122	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: DNA	
A;Residues: 1-780 <TAK>	
A;Cross-references: UNIPROT:P97040; UNIPARC:UPI00000BCBC4; EMBL:AB010203; NID:G2780763; I	
A;Experimental source: strain Ictero No.1; substrain icterohaemorrhagiae	
C;Superfamily: Leptospira interrogans hypothetical 88.9K protein	
Query Match 89.3%; Score 25; DB 2; Length 780;	
Best Local Similarity 55.6%; Pred. No. 3.3e+02;	
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 VRDQIXXXL 9 :
Db	734 IRDQIAEFL 742
RESULT 22	
T21910 hypothetical protein F37B12.4 - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C;Accession: T21910	
R;Thomas, K.	
submitted to the EMBL Data Library, September 1995	
A;Reference number: Z19486	
A;Accession: T21910	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: DNA	
A;Residues: 1-1430 <WIL>	
A;Cross-references: UNIPROT:O01049; UNIPARC:UPI0000080F77; EMBL:Z54218; PIDN:CAA90957.1;	
A;Experimental source: clone F37B12	
C;Genetics:	
A;Gene: CESP:F37B12.4	
A;Map position: 2	

A;Introns: 319/2; 666/1; 714/2; 1011/2; 1065/3; 1116/2; 1189/3; 1214/3; 1283/2

Query Match 89.3%; Score 25; DB 2; Length 1430;
Best Local Similarity 55.6%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 135 IRDQIKSLL 143

RESULT 23
T22812
hypothetical protein ZC116.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22812; T27494
R;Burton, J.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19618
A;Accession: T22812
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3871 <WIL>
A;Cross-references: UNIPROT:Q20911; UNIPARC:UPI00000821EA; EMBL:Z74473; PIDN:CAA98952.1;
A;Experimental source: clone F56H9
R;Smye, R.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z20376
A;Accession: T27494
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3871 <WI2>
A;Cross-references: UNIPARC:UPI00000821EA; EMBL:Z74046; PIDN:CAA98557.1; GSPDB:GN00023;
A;Experimental source: clone ZC116
C;Genetics:
A;Gene: CESP:ZC116.3
A;Map position: 5
A;Introns: 29/3; 66/2; 244/1; 359/3; 422/1; 549/1; 572/1; 728/1; 820/1; 889/3; 1189/1; 1
2763/1; 2809/1; 2906/3; 3051/3; 3176/1; 3254/1; 3341/2; 3536/1; 3565/2; 3582/3; 3609/1;
C;Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology

Query Match 89.3%; Score 25; DB 2; Length 3871;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 1714 VRDQIGFVL 1722

RESULT 24
G82442
Sui1 family protein VCA0570 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82442
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <HEI>
A;Cross-references: UNIPROT:Q9KM19; UNIPARC:UPI00000C3588; GB:AE004388; GB:AE003853; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0570
A;Map position: 2
C;Superfamily: translation initiation factor SUI

Query Match 85.7%; Score 24; DB 2; Length 103;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 83 VRDQLKTLL 91

RESULT 25
A87425
hypothetical protein CC1416 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87425
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <STO>
A;Cross-references: UNIPROT:Q9A8D8; UNIPARC:UPI00000C739E; GB:AE005673; NID:g13422777; P
C;Genetics:
A;Gene: CC1416

Query Match 85.7%; Score 24; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
:|||||
Db 31 VRDQI 35

RESULT 26
T39825
hypothetical protein SPBC19F8.02 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39825
R;Beck, A.; Reinhardt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21883
A;Accession: T39825
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-166 <BEC>
A;Cross-references: UNIPROT:O60166; UNIPARC:UPI00000C67AF; EMBL:AL0233594; NID:e12933398;
A;Experimental source: strain 972h-; cosmid c19F8
C;Genetics:
A;Gene: SPDB:SPBC19F8.02
A;Map position: 2

Query Match 85.7%; Score 24; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
:|||||
Db 162 VRDQI 166

RESULT 27
E83662
hypothetical protein BH0101 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: E83662

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: E83662
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <STO>
A;Cross-references: UNIPROT:Q9JWP8; UNIPARC:UPI00000C492E; GB:AP001507; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0101
C;Superfamily: Bacillus subtilis hypothetical protein yacH

Query Match 85.7%; Score 24; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|
|
|
|
Db 161 VRDQI 165

RESULT 28
S66113
hypothetical protein.yacH - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66113; A69741
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom
A;Reference number: S65967; MUID:96051385; PMID:7584024
A;Accession: S66113
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-185 <GA>
A;Cross-references: UNIPROT:P37569; UNIPARC:UPI000005FE01; EMBL:D26185; NID:g467326; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Sadale, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69741

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

QY 1 VRDQI 5
|
|
|
|
Db 164 VRDQI 168

hypothetical protein.yacH

Query Match 85.7%; Score 24; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 29
S39401
hydrogenase maturation factor hupD [similarity] - Bradyrhizobium japonicum
C;Species: Bradyrhizobium japonicum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S39401
R;van Soom, C.; Browaeys, J.; Verreth, C.; Vanderleyden, J.
J. Mol. Biol. 234, 508-512, 1993
A;Title: Nucleotide sequence analysis of four genes, hupC, hupD, hupF and hupG, downstre
A;Reference number: S39400; MUID:94047099; PMID:8230232
A;Accession: S39401
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-193 <VAN>
A;Cross-references: UNIPROT:Q45251; UNIPARC:UPI000016E70D; EMBL:Z21948; NID:g311536; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
C;Superfamily: [NiFe]-hydrogenase maturation protease

Query Match 85.7%; Score 24; DB 1; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|
|
|
|
Db 142 VRDQI 146

RESULT 30
B82601

transcription regulator AcrR family XF2085 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: B82601
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82601
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-205 <SIM>
A;Cross-references: UNIPROT:Q9PBQ5; UNIPARC:UPI00000C2924; GB:AE004024; GB:AE003849; NID
A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2085

Query Match 85.7%; Score 24; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|
|
|
|
Db 19 VRDQI 23

RESULT 31

F81658
type III secretion translocase sctL TC0850 [imported] - Chlamydia muridarum (strain Nigg
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81658
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <TET>
A;Cross-references: UNIPROT:Q9PJH9; UNIPARC:UPI0000057AA8; GB:AE002351; GB:AE002160; NID
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0850

Query Match 85.7%; Score 24; DB 2; Length 235;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 98 MRDQIKSSL 106

RESULT 32

E70104
1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) (plsc) - Lyme disease spiroch
N;Alternate names: 1-acyl-sn-glycerol-3-phosphate acetyltransferase
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 16-Aug-2004
C;Accession: E70104; I40293
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: E70104
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-250 <KLE>
A;Cross-references: UNIPROT:Q59188; UNIPARC:UPI0000057312; GB:AE001117; GB:AE000783; NID
A;Experimental source: strain B31
R;Ojaimi, C.; Davidson, B.E.; Saint Giron, I.; Old, I.G.
Microbiology 140, 2931-2940, 1994
A;Title: Conservation of gene arrangement and an unusual organization of rRNA genes in b
A;Reference number: I40241; MUID:95111614; PMID:7812434
A;Accession: I40293
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 116-250 <RES>
A;Cross-references: UNIPARC:UPI000016E69A; GB:L32861; NID:g520779; PIDN:AAC41407.1; PID:
C;Genetics:
A;Gene: plsc
C;Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase
C;Keywords: acyltransferase; coenzyme A

Query Match 85.7%; Score 24; DB 2; Length 250;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 237 IRDQIVKKL 245

RESULT 33
AE0127

Conserved hypothetical protein YPO1038 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE0127
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-257 <KUR>
A;Cross-references: UNIPROT:Q8ZH72; UNIPARC:UPI00000DCCEF; GB:AL590842; PIDN:CAC89880.1;
C;Genetics:
A;Gene: YPO1038
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
:||||
Db 39 VRDQI 43

RESULT 34

A69965
ketoacyl reductase homolog yqjQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C;Accession: A69965
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertei
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69965
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-259 <KUN>
A;Cross-references: UNIPROT:P54554; UNIPARC:UPI00000606B5; GB:Z99116; GB:AL009126; NID:
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqjQ
C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
F;7-183/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 85.7%; Score 24; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
:||||
Db 74 VRDQI 78

RESULT 35

AG0861
probable RNA pseudouridylyate synthase ygcB [imported] - Salmonella enterica subsp. ente

C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0861
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <PAR>
A;Cross-references: UNIPARC:UPI000005A3BD; GB:AL513382; PIDN:CAD06078.1; PID:G16504044;
C;Genetics:
A;Gene: YgcB
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 39 VRDQI 43

RESULT 36
C91085
hypothetical protein ECs3651 [imported] - Escherichia coli (strain O157:H7, substrain R)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C91085
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C91085
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <HAY>
A;Cross-references: UNIPROT:Q8X6T6; UNIPARC:UPI00000D0B23; GB:BA000007; PIDN:BAB37074.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs3651
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 39 VRDQI 43

RESULT 37
E85930
hypothetical protein Z4107 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85930
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85930
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-260 <STO>
A;Cross-references: UNIPROT:Q8X6T6; UNIPARC:UPI00000D0B23; GB:AE005174; NID:G12517267; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4107
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 39 VRDQI 43

RESULT 38
C65061
hypothetical protein b2791 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C65061
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C65061
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-260 <BLAT>
A;Cross-references: UNIPROT:Q46918; UNIPARC:UPI000013BED9; GB:AE000363; GB:U00096; NID:G
A;Experimental source: strain K-12, substrain MGL655
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 39 VRDQI 43

RESULT 39
T22929
hypothetical protein F58G1.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22929
R;Smye, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19639
A;Accession: T22929
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-302 <WIL>
A;Cross-references: UNIPROT:O62271; UNIPARC:UPI00000801FB; EMBL:Z81556; PIDN:CAB04520.1;
A;Experimental source: clone F58G1
C;Genetics:
A;Gene: CESP:F58G1.5
A;Map position: 2
A;Introns: 89/2; 127/3; 162/1; 231/3; 262/3
C;Superfamily: Caenorhabditis elegans hypothetical protein C09E8.2

Query Match 85.7%; Score 24; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 37 VRDQI 41

RESULT 40
CPCYB
carboxypeptidase B (EC 3.4.17.2) - broad-fingered crayfish
C;Species: Astacus astacus, Astacus fluviatilis (broad-fingered crayfish)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A05141; A14178
R;Titani, K.; Ericsson, L.H.; Kumar, S.; Jakob, F.; Neurath, H.; Zwilling, R.
Biochemistry 23, 1245-1250, 1984
A;Title: Amino acid sequence of crayfish (Astacus fluviatilis) carboxypeptidase B.
A;Reference number: A90482
A;Accession: A05141
A;Molecule type: protein
A;Residues: 1-303 <TIT>
A;Cross-references: UNIPROT:P04069; UNIPARC:UPI00001271CA
C;Superfamily: carboxypeptidase
C;Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein digestion; zinc
F;241,264/Active site: Tyr, Glu #status predicted

Query Match 85.7%; Score 24; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 171 VRDQI 175

RESULT 41
H87447
conserved hypothetical protein CC1601 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: H87447
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: H87447
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-315 <STO>
A;Cross-references: UNIPROT:Q9A7W8; UNIPARC:UPI00000C7439; GB:AE005673; NID:gl3422998; P
C;Genetics:
A;Gene: CC1601

Query Match 85.7%; Score 24; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 94 VRDQI 98

RESULT 42
C87271
general secretion pathway protein K [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87271
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87271
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>

A;Cross-references: UNIPROT:Q9ABP6; UNIPARC:UPI00000C6F67; GB:AE005673; NID:gl3421299; P
C;Genetics:
A;Gene: CC0180
C;Superfamily: outK protein

Query Match 85.7%; Score 24; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
| | | | |
Db 275 VRDQI 279

RESULT 43
T29369
hypothetical protein ZC404.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29369
R;Bentley, D.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid ZC404.
A;Reference number: Z20614
A;Accession: T29369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-351 <BEN>
A;Cross-references: UNIPROT:Q23287; UNIPARC:UPI0000001CFD; EMBL:U55363; PIDN:AAA97963.1;
A;Experimental source: strain Bristol N2; clone ZC404
C;Genetics:
A;Gene: CESP:ZC404.8
A;Map position: 5
A;Introns: 17/2; 52/2; 73/2; 312/1

Query Match 85.7%; Score 24; DB 2; Length 351;
Best Local Similarity 55.6%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
: | | | | |
Db 267 LRDQITALL 275

RESULT 44
T04917
hypothetical protein T10114.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04917
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.x
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15389
A;Accession: T04917
A;Molecule type: DNA
A;Residues: 1-369 <BEV>
A;Cross-references: UNIPROT:O49641; UNIPARC:UPI00000A4B3A; EMBL:AL021712
A;Experimental source: cultivar Columbia; BAC clone T10114
C;Genetics:
A;Map position: 4
A;Introns: 183/2; 220/3; 278/1
A;Note: T10114.190

Query Match 85.7%; Score 24; DB 2; Length 369;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
: | | | | |
Db 48 IRDQINILL 56

RESULT 45

S45107
hypothetical protein 1 - Erwinia carotovora
C;Species: Erwinia carotovora
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S45107
R;Golby, P.; Jones, S.E.; Stephens, S.; Reeves, P.J.; Bycroft, B.; Stewart, G.; Williams submitted to the EMBL Data Library, May 1994
A;Description: Global regulation of Erwinia carotovora exoenzyme virulence factors: mult
A;Reference number: S45107
A;Accession: S45107
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <GOL>
A;Cross-references: UNIPROT:Q47417; UNIPARC:UPI000013BEDA; EMBL:X79474; NID:g496597; PID

Query Match 85.7%; Score 24; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|||
Db 153 VRDQI 157

RESULT 46
T32710
hypothetical protein T22D1.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32710
R;Geisel, C.; Bradshaw, H.; Hawkins, M.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid T22D1.
A;Reference number: Z21211
A;Accession: T32710
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-448 <GEI>
A;Cross-references: UNIPROT:Q9GZH2; UNIPARC:UPI000007D4F4; EMBL:AF039052; PIDN:AAB94278.
A;Experimental source: strain Bristol N2; clone T22D1
C;Genetics:
A;Gene: CESP:T22D1.10
A;Map position: 4
A;Introns: 14/3; 60/3; 103/3; 196/1; 326/2; 377/2
C;Superfamily: conserved hypothetical protein YDR190c

Query Match 85.7%; Score 24; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|||
Db 268 VRDQI 272

RESULT 47
S61029
hypothetical protein YPL235w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein P1060
C;Species: Saccharomyces cerevisiae
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S61029; S61699; S65260; S65264
R;Pohl, T.M.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61010
A;Accession: S61029
A;Molecule type: DNA
A;Residues: 1-471 <POH>
A;Cross-references: UNIPROT:Q12464; UNIPARC:UPI0000052E06; EMBL:Z67751; NID:g1061234; PI
R;Urrestarazu, L.A.
submitted to the EMBL Data Library, December 1995
A;Reference number: S61699
A;Accession: S61699

A;Molecule type: DNA
A;Residues: 1-471 <URR>
A;Cross-references: UNIPARC:UPI0000052E06; EMBL:X94561; NID:g1181252; PIDN:CAA64252.1; P
R;Urrestarazu, L.A.; Visser, S.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65251
A;Accession: S65260
A;Molecule type: DNA
A;Residues: 1-471 <URF>
A;Cross-references: UNIPARC:UPI0000052E06; EMBL:Z73591; NID:g1370485; PIDN:CAA97952.1; P
A;Experimental source: strain S288C (AB972)
R;Pohl, T.M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64899
A;Accession: S65264
A;Molecule type: DNA
A;Residues: 1-471 <POW>
A;Cross-references: UNIPARC:UPI0000052E06; EMBL:Z73591; NID:g1370485; PIDN:CAA97952.1; P
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Gene: SGD:RVB2
A;Cross-references: SGD:S0006156
A;Map position: 16L
C;Superfamily: conserved hypothetical protein YDR190c

Query Match 85.7%; Score 24; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
|||
Db 269 VRDQI 273

RESULT 48
S64837
hypothetical protein YLR015w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L1583
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C;Accession: S64837; S59274; S48527
R;Vandenbol, M.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64742
A;Accession: S64837
A;Molecule type: DNA
A;Residues: 1-505 <VAN>
A;Cross-references: UNIPROT:P43132; UNIPARC:UPI0000052FF0; EMBL:Z73187; NID:g1360313; PI
A;Experimental source: strain S288C
R;Saville, S.P.; Atkinson, S.; Jamieson, L.; Pocklington, M.J.; Orr, E.
submitted to the EMBL Data Library, August 1995
A;Description: A 7.8kb fragment from chromosome XII of Saccharomyces cerevisiae does not
A;Reference number: S59270
A;Accession: S59274
A;Molecule type: DNA
A;Residues: 1-464,'ISLL',468 <SAV>
A;Cross-references: UNIPARC:UPI000006AFA4; EMBL:X90564; NID:g975221; PIDN:CAA62158.1; PI
A;Experimental source: strain S288C
R;Levin, D.E.; Stevenson, W.D.
submitted to the EMBL Data Library, July 1994
A;Description: The S.cerevisiae PKC2 does not exist in the yeast genome.
A;Reference number: S48527
A;Accession: S48527
A;Molecule type: DNA
A;Residues: 1-472 <LEV>
A;Cross-references: UNIPARC:UPI0000179EB1; EMBL:L34405
C;Genetics:
A;Gene: SGD:BRE2
A;Cross-references: SGD:S0004005
A;Map position: 12R

Query Match 85.7%; Score 24; DB 2; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Search completed: May 12, 2006, 10:52:38
Job time : 19.0231 secs

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	309	VRDQI 313							
RESULT 49									
PQ0646									
ferredoxin-nitrite reductase (EC 1.7.7.1) - Synechococcus sp. (strain PCC 7942)									
C;Species: Synechococcus sp.									
A;Variety: PCC 7942									
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004									
C;Accession: S33530; PQ0646; S32143									
R;Luque, I.; Flores, E.; Herrero, A.									
Plant Mol. Biol. 21, 1201-1205, 1993									
A;Title: Nitrite reductase gene from Synechococcus sp. PCC 7942: homology between cyanob									
A;Reference number: S33530; MUID:93257637; PMID:8490140									
A;Accession: S33530									
A;Molecule type: DNA									
A;Residues: 1-512 <LUQ>									
A;Cross-references: UNIPROT:P39661; UNIPARC:UPI00001301CB; EMBL:X67680; NID:g288053; PID									
R;Omata, T.									
Plant Cell Physiol. 32, 151-157, 1991									
A;Title: Cloning and characterization of the nrtA gene that encodes a 45-kDa protein inv									
A;Reference number: PQ0646									
A;Accession: PQ0646									
A;Molecule type: DNA									
A;Residues: 468-512 <OMA>									
A;Cross-references: UNIPARC:UPI0000172138; DDBJ:D00677									
C;Genetics:									
A;Gene: nir									
C;Superfamily: sulfite/ferredoxin-nitrite reductase									
C;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein; oxidoreductase									
F;396,402,437,441/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted									
Query Match 85.7%; Score 24; DB 1; Length 512;									
Best Local Similarity 100.0%; Pred. No. 3.7e+02;									
Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	26	VRDQI 30							
RESULT 50									
G83558									
hypothetical protein PA0696 [imported] - Pseudomonas aeruginosa (strain PA01)									
C;Species: Pseudomonas aeruginosa									
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004									
C;Accession: G83558									
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B									
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,									
.; Lory, S.; Olson, M.V.									
Nature 406, 959-964, 2000									
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho									
A;Reference number: A82950; MUID:20437337; PMID:10984043									
A;Accession: G83558									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-568 <STO>									
A;Cross-references: UNIPROT:Q915N0; UNIPARC:UPI00000C5107; GB:AE004505; GB:AE004091; NID									
A;Experimental source: strain PA01									
C;Genetics:									
A;Gene: PA0696									
Query Match 85.7%; Score 24; DB 2; Length 568;									
Best Local Similarity 100.0%; Pred. No. 4.2e+02;									
Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	96	VRDQI 100							

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 63.6923 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Match	Length	DB ID	Description
		Match	Length				
1	27	96.4	232	2	Q6MGX7	BDEBA	Q6ngx7 bdellovibri
2	27	96.4	455	2	Q7QAB5	ANOGA	Q7qab5 anopheles g
3	27	96.4	543	2	Q9ZSH5	ARATH	Q9zsh5 arabidopsis
4	27	96.4	1113	2	Q9M180	ARATH	Q9m180 arabidopsis
5	27	96.4	1119	2	Q9LK04	ARATH	Q9lk04 arabidopsis
6	27	96.4	1132	2	Q9S9Z6	ARATH	Q9s9z6 arabidopsis
7	27	96.4	1198	2	Q9C6N9	ARATH	Q9c6n9 arabidopsis
8	27	96.4	1201	2	Q9C607	ARATH	Q9c607 arabidopsis
9	27	96.4	1305	2	Q9ZU21	ARATH	Q9zu21 arabidopsis
10	27	96.4	1444	2	Q9LTG8	ARATH	Q9ltg8 arabidopsis
11	27	96.4	1444	2	Q9SFY0	ARATH	Q9sfy0 arabidopsis
12	27	96.4	1472	2	Q9SL11	ARATH	Q9sl11 arabidopsis
13	27	96.4	1756	2	O80975	ARATH	O80975 arabidopsis
14	26	92.9	211	2	O65247	ARATH	O65247 arabidopsis
15	26	92.9	218	2	Q7XJK7	ARATH	Q7xjk7 arabidopsis
16	26	92.9	272	2	Q5WH83	BACSK	Q5wh83 bacillus cl
17	26	92.9	334	2	Q7NKK0	GLOVI	Q7nkk0 gloeobacter
18	26	92.9	357	2	Q8Z110	SALTI	Q8z110 salmonella
19	26	92.9	486	2	O8MYV0	DROME	O8myv0 drosophila
20	26	92.9	486	2	Q9VW68	DROME	Q9vw68 drosophila
21	26	92.9	516	2	Q4PCC1	USTMA	Q4pcc1 ustilago ma
22	26	92.9	594	2	Q4Q7Y6	LEIMA	Q4q7y6 leishmania
23	26	92.9	642	2	Q89S89	BRAJA	Q89s89 bradyrhizob
24	26	92.9	743	2	Q5LX00	SILPO	Q5lxx0 silicibacte
25	26	92.9	896	2	Q7UYC9	RHOBA	Q7uyc9 rhodopirell
26	26	92.9	1084	2	Q65353	NPVAC	Q65353 autographa
27	26	92.9	1105	2	Q4P3H5	USTMA	Q4p3h5 ustilago ma
28	26	92.9	1191	2	Q4I1R0	GIBZE	Q4ilr0 gibberella
29	26	92.9	1312	2	Q9M0W7	ARATH	Q9m0w7 arabidopsis
30	26	92.9	1407	2	Q61Z65	CAEBR	Q61z65 caenorhabdi
31	26	92.9	2240	2	Q51WD5	MAGGR	Q51wd5 magnaporth

32	25	89.3	89	2	Q76YV2	9CAUD	Q76yv2 bacterioph
33	25	89.3	102	2	Q6RXD6	HCMV	Q6rxd6 human cytom
34	25	89.3	115	1	UL101	HCMVA	P16826 human cytom
35	25	89.3	130	2	Q58J60	9ACTO	Q58j60 streptomyce
36	25	89.3	139	2	Q4XMZ2	PLACH	Q4xmz2 plasmodium
37	25	89.3	193	2	Q9I6T7	PSEAE	Q9i6t7 pseudomonas
38	25	89.3	225	2	Q65L06	BACLD	Q65l06 bacillus li
39	25	89.3	229	2	Q6VT06	9VIRU	Q6vt06 vibrio para
40	25	89.3	243	2	P90912	CAEEL	P90912 caenorhabdi
41	25	89.3	250	1	EVG1	DROME	Q9vss7 drosophila
42	25	89.3	250	2	Q8MZ83	DROME	Q8mz83 drosophila
43	25	89.3	291	2	Q72MR9	LEPIC	Q72mr9 leptospira
44	25	89.3	291	2	Q8EZE2	LEPIN	Q8eze2 leptospira
45	25	89.3	296	2	Q8G822	BIFLO	Q8g822 bifidobacte
46	25	89.3	327	1	PERN	ARMRU	Q42517 armoracia r
47	25	89.3	358	2	Q61Y40	CAEBR	Q61y40 caenorhabdi
48	25	89.3	422	2	Q4WFK8	ASPFU	Q4wfk8 aspergillus
49	25	89.3	437	2	Q8D8W3	VIBVU	Q8d8w3 vibrio vuln
50	25	89.3	437	2	Q7MLK9	VIBVY	Q7mlk9 vibrio vuln
51	25	89.3	452	2	Q9LJ56	ARATH	Q9lj56 arabidopsis
52	25	89.3	495	2	Q6D280	ERWCT	Q6d280 erwinia car
53	25	89.3	504	2	Q6FN90	CANGA	Q6fn90 candida gla
54	25	89.3	507	1	CDC7	YEAST	P06243 saccharomyc
55	25	89.3	516	2	Q5YEL8	SOLTU	Q5yel8 solanum tub
56	25	89.3	522	2	Q525T3	MAGGR	Q525t3 magnaporth
57	25	89.3	530	1	GUAA	BORPA	Q7wk13 bordetella
58	25	89.3	530	1	GUAA	BORPE	Q7wav6 bordetella
59	25	89.3	530	1	GUAA	BORPE	Q7vvm4 bordetella
60	25	89.3	551	1	G6PI	SYNPX	Q7u6t0 synchococc
61	25	89.3	587	2	Q8PYB9	METMA	Q8pyb9 methanosarc
62	25	89.3	610	2	Q9KRY4	VIBCH	Q9kry4 vibrio chol
63	25	89.3	619	2	Q7DLA2	LEPIN	Q7dla2 leptospira
64	25	89.3	621	2	Q8U9J0	AGRT5	Q8u9j0 agrobacteri
65	25	89.3	624	2	Q5WVJ4	LEGPL	Q5wvj4 legionella
66	25	89.3	624	2	Q5X455	LEGPA	Q5x455 legionella
67	25	89.3	624	2	Q5ZUE3	LEGPH	Q5zue3 legionella
68	25	89.3	635	2	Q4H812	9DEIO	Q4h812 deinococcus
69	25	89.3	642	2	Q4YN59	PLABE	Q4yn59 plasmodium
70	25	89.3	642	2	Q7CT69	AGRT5	Q7ct69 agrobacteri
71	25	89.3	686	1	RECG	STAAN	P64324 staphylococ
72	25	89.3	686	1	RECG	STAAN	P64325 staphylococ
73	25	89.3	686	1	RECG	STAAU	O50581 staphylococ
74	25	89.3	686	1	RECG	STAAW	Q8nx11 staphylococ
75	25	89.3	686	2	Q6G9Y6	STAAS	Q6g9y6 staphylococ
76	25	89.3	686	2	Q6GHK8	STAAR	Q6ghk8 staphylococ
77	25	89.3	686	2	Q5HGK6	STAAC	Q5hgk6 staphylococ
78	25	89.3	737	2	Q5H511	XANOR	Q5h511 xanthomonas
79	25	89.3	737	2	Q8PQ22	XANAC	Q8pq22 xanthomonas
80	25	89.3	747	2	Q4UZD4	XANCP	Q4uzd4 xanthomonas
81	25	89.3	747	2	Q8PD49	XANCP	Q8pd49 xanthomonas
82	25	89.3	771	2	Q72AX1	DESVH	Q72ax1 desulfovibr
83	25	89.3	780	2	P97040	LEPIN	P97040 leptospira
84	25	89.3	784	2	Q72VZ2	LEPIC	Q72vz2 leptospira
85	25	89.3	784	2	Q8F9L9	LEPIN	Q8f9l9 leptospira
86	25	89.3	793	2	Q5AS07	EMENTI	Q5as07 aspergillus
87	25	89.3	827	2	Q7UM67	RHOBA	Q7um67 rhodopirell
88	25	89.3	836	2	Q4FEY1	9VIRU	Q4fey1 cotesia plu
89	25	89.3	853	2	Q7RDQ3	PLAYO	Q7rdq3 plasmodium
90	25	89.3	960	2	Q7S6I2	NEUCR	Q7s6i2 neurospora
91	25	89.3	965	2	Q6SLD5	COCHE	Q6sld5 cochliobolu
92	25	89.3	988	2	P93661	CHEAL	P93661 chenopodium
93	25	89.3	1073	2	Q82QD5	STRAW	Q82qd5 streptomyce
94	25	89.3	1779	2	Q5CUN7	CRYPV	Q5cun7 cryptospori
95	25	89.3	1780	2	Q5CPD0	CRYHO	Q5cpd0 cryptospori
96	25	89.3	2183	2	Q6C4J0	YARLI	Q6c4j0 yarrowia li
97	25	89.3	2920	2	P90865	CABEL	P90865 caenorhabdi
98	25	89.3	2923	2	Q7JLZ7	CABEL	Q7j1z7 caenorhabdi
99	25	89.3	3251	2	Q5K8V2	CRYNE	Q5kbv2 cryptococcu
100	25	89.3	3342	2	Q55NG7	CRYNE	Q55ng7 cryptococcu
101	25	89.3	3871	2	Q20911	CABEL	Q20911 caenorhabdi
102	24	85.7	26	2	Q6V7G4	VIBCH	Q6v7g4 vibrio chol
103	24	85.7	39	2	Q4TIC1	TETNG	Q4tic1 tetraodon n
104	24	85.7	50	2	Q4SUG6	TETNG	Q4sug6 tetraodon n

105	24	85.7	70	2	Q853Q9_9CAUD	Q853Q9_mycobacteri
106	24	85.7	74	2	Q58L62_BRACM	Q58L62_brassica ca
107	24	85.7	84	2	Q878D6_STRP3	Q878d6 streptococc
108	24	85.7	84	2	Q5XDG1_STRP6	Q5xdg1 streptococc
109	24	85.7	84	2	Q9A140_STRPY	Q9a140 streptococc
110	24	85.7	84	2	Q8P264_STRP8	Q8p264 streptococc
111	24	85.7	88	2	Q95IQ5_BOVIN	Q95iq5 bos taurus
112	24	85.7	95	2	Q4KGC5_PSEF5	Q4kgc5 pseudomonas
113	24	85.7	103	2	Q9KM19_VIBCH	Q9km19 vibrio chol
114	24	85.7	104	2	Q6N279_RHOPA	Q6n279 rhodopseudo
115	24	85.7	109	2	Q9A8D8_CAUCR	Q9a8d8 caulobacter
116	24	85.7	129	2	Q8NTP3_CORGL	Q8ntp3 corynebacte
117	24	85.7	130	2	Q9LMS4_ARATH	Q9lms4 arabidopsis
118	24	85.7	139	2	Q7V4I1_PROMM	Q7v4i1 prochloroco
119	24	85.7	139	2	Q7U457_SYNPM	Q7u457 synchococc
120	24	85.7	141	2	Q890G2_LACPL	Q890g2 lactobacill
121	24	85.7	142	2	Q86DC5_CAEEL	Q86dc5 caenorhabdi
122	24	85.7	151	2	Q99QJ0_STRCO	Q99qj0 streptomyce
123	24	85.7	152	2	Q64G27_CLOIN	Q64g27 clostridium
124	24	85.7	157	2	Q6AM17_DESPS	Q6am17 desulfotale
125	24	85.7	166	2	Q60166_SCHPO	Q60166 schizosacch
126	24	85.7	168	2	Q5Y4A6_CARPA	Q5y4a6 carica papa
127	24	85.7	175	2	Q86HU6_DICDI	Q86hu6 dictyosteli
128	24	85.7	175	2	Q8BI02_MOUSE	Q8bi02 mus musculu
129	24	85.7	177	2	Q9JWP8_BACHD	Q9jwp8 bacillus ha
130	24	85.7	182	2	Q5L438_GEOKA	Q5l438 geobacillus
131	24	85.7	183	2	Q7SCG4_NEUCR	Q7scg4 neurospora
132	24	85.7	185	1	YACH_BACSU	P37569 bacillus su
133	24	85.7	192	1	HUPD_BRAJA	Q45251 bradyrhizob
134	24	85.7	204	2	Q88HQ3_PSEPK	Q88hq3 pseudomonas
135	24	85.7	204	2	Q8XUI1_RALSO	Q8xui1 ralstonia s
136	24	85.7	205	2	Q4LYE5_9BURK	Q4lye5 burkholderi
137	24	85.7	205	2	Q9PBQ5_XYLFA	Q9pbq5 xylella fas
138	24	85.7	205	2	Q8PPQ91_XANAC	Q8ppq91 xanthomonas
139	24	85.7	206	2	Q7Q261_ANOGA	Q7q261 anopheles g
140	24	85.7	207	2	Q4UZK6_XANCP	Q4uzk6 xanthomonas
141	24	85.7	207	2	Q8PDC0_XANCP	Q8pdc0 xanthomonas
142	24	85.7	209	2	Q5N3V5_SYNPM	Q5n3v5 conioophora
143	24	85.7	211	2	Q9XLN0_9HOMO	Q9n3v5 synchococc
144	24	85.7	212	2	Q9XLN0_9HOMO	Q9xln0 paxillus fi
145	24	85.7	212	2	Q9XLN1_PAXIN	Q9xln1 paxillus in
146	24	85.7	212	2	Q9XLN7_9HOMO	Q9xln7 phaegyropo
147	24	85.7	212	2	Q9XLN8_9HOMO	Q9xln8 boletinelu
148	24	85.7	212	2	Q9XLN9_9HOMO	Q9xln9 gyrodon liv
149	24	85.7	212	2	Q9XLP1_9HOMO	Q9xlp1 scleroderma
150	24	85.7	216	2	Q79606_BOLED	Q79606 boletus edu
151	24	85.7	219	2	Q9XLM9_9HOMO	Q9xlm9 chalciporus
152	24	85.7	221	2	Q79607_9HOMO	Q79607 boletus pal
153	24	85.7	221	2	Q79608_9HOMO	Q79608 xerocomus c
154	24	85.7	221	2	Q9XLM7_9HOMO	Q9xlm7 phylloporus
155	24	85.7	221	2	Q9XLM8_9HOMO	Q9xlm8 chamonixia
156	24	85.7	221	2	Q554V9_DICDI	Q554v9 dictyosteli
157	24	85.7	221	2	Q4W494_AERSA	Q4w494 aeromonas s
158	24	85.7	221	2	Q699P2_AERHY	Q699p2 aeromonas h
159	24	85.7	225	2	Q51L66_MAGGR	Q51l66 magnaporthe
160	24	85.7	227	2	Q6AG89_LEIXX	Q6ag89 leifsonia x
161	24	85.7	229	2	Q5CJ16_CRYHO	Q5cji6 cryptospori
162	24	85.7	229	2	Q88A14_PSESM	Q88a14 pseudomonas
163	24	85.7	231	2	Q8NCV3_HUMAN	Q8ncv3 homo sapien
164	24	85.7	235	2	Q9PJH9_CHLMU	Q9pjh9 chlamydia m
165	24	85.7	240	2	Q6E6F0_ANTLO	Q6e6f0 antonospora
166	24	85.7	245	2	Q662X5_BORGA	Q662x5 borrelia ga
167	24	85.7	246	2	Q7Q255_ANOGA	Q7q255 anopheles g
168	24	85.7	247	2	Q56BU1_9CAUD	Q56bu1 enterobacte
169	24	85.7	250	1	PLSC_BORBU	Q59l88 borrelia bu
170	24	85.7	252	1	SURE_RHILO	Q98lcn rhizobium l
171	24	85.7	254	2	Q8HN57_CRYNE	Q8hn57 cryptococcu
172	24	85.7	254	2	Q6J9U3_CRYNV	Q6j9u3 cryptococcu
173	24	85.7	254	2	Q8NMU6_CORGL	Q8nmu6 corynebacte
174	24	85.7	255	2	Q4K3B2_PSEF5	Q4k3b2 pseudomonas
175	24	85.7	256	2	Q5JST5_HUMAN	Q5jst5 homo sapien
176	24	85.7	257	1	TRUC_YERPE	Q8zh72 yersinia pe
177	24	85.7	257	2	Q667I3_YERPS	Q667i3 yersinia ps

178	24	85.7	259	1	YQJQ_BACSU	P54554 bacillus su
179	24	85.7	259	2	Q4ZT89_PSESY	Q4zt89 pseudomonas
180	24	85.7	260	1	TRUC_ECO57	Q8x6t6 escherichia
181	24	85.7	260	1	TRUC_ECOL6	Q8fef9 escherichia
182	24	85.7	260	1	TRUC_ECOLI	Q469i8 escherichia
183	24	85.7	260	1	TRUC_SALTI	Q8z439 salmonella
184	24	85.7	260	1	TRUC_SALTY	Q8zmd5 salmonella
185	24	85.7	260	2	Q57KF2_SALCH	Q57kf2 salmonella
186	24	85.7	260	2	Q5PEK3_SALPA	Q5pek3 salmonella
187	24	85.7	261	2	Q6D8E9_ERWCT	Q6d8e9 erwinia car
188	24	85.7	262	2	Q5CU68_CRYPV	Q5cu68 cryptospori
189	24	85.7	269	2	Q8CFH8_MOUSE	Q8cfh8 mus musculu
190	24	85.7	270	2	Q9LM35_ARATH	Q9lm35 arabidopsis
191	24	85.7	277	2	Q85SZ9_CRYNV	Q85sz9 cryptococcu
192	24	85.7	280	2	Q61R33_CAEBR	Q61r33 caenorhabdi
193	24	85.7	285	2	Q4QPJ5_HAEI8	Q4qpj5 haemophilus
194	24	85.7	287	2	Q89L16_BRAJA	Q89l16 bradyrhizob
195	24	85.7	292	2	Q6M2Z7_CORGL	Q6m2z7 corynebacte
196	24	85.7	294	2	Q89HJ2_BRAJA	Q89hj2 bradyrhizob
197	24	85.7	295	2	Q4N167_THEPA	Q4n167 thelleria p
198	24	85.7	295	2	Q5FM19_LACAC	Q5fm19 lactobacill
199	24	85.7	296	2	Q74KG0_LACJO	Q74kg0 lactobacill
200	24	85.7	302	2	Q62271_CAEEL	Q62271 caenorhabdi
201	24	85.7	302	2	Q87H02_VIBPA	Q87h02 vibrio para
202	24	85.7	303	1	CBPB_ASTFL	P04069 astacus flu
203	24	85.7	306	2	Q8IN95_DROME	Q8in95 drosophila
204	24	85.7	307	2	Q75H63_ORYSA	Q75h63 oryza sativ
205	24	85.7	315	2	Q9A7W8_CAUCR	Q9a7w8 caulobacter
206	24	85.7	316	1	CCSA_ADICA	Q85fh6 adiantum ca
207	24	85.7	316	2	Q6FAS4_ACIAD	Q6fas4 acinetobact
208	24	85.7	316	2	Q9EQQ5_MOUSE	Q9eqg5 mus musculu
209	24	85.7	317	1	T2N1_NOCAE	P50187 nocardia ae
210	24	85.7	319	2	Q4LUR9_9BURK	Q4lur9 burkholderi
211	24	85.7	324	2	Q9ABP6_CAUCR	Q9abp6 caulobacter
212	24	85.7	325	2	Q50WB0_ENTHI	Q50wb0 entamoeba h
213	24	85.7	328	2	Q8KPM8_SACER	Q8kpw8 saccharopol
214	24	85.7	333	2	Q8GW82_ARATH	Q8gw82 arabidopsis
215	24	85.7	338	1	RFC3_ARXAD	Q74l11 arxula aden
216	24	85.7	338	2	Q5LNQ8_SILPO	Q5lnq8 silicibacte
217	24	85.7	340	2	Q4H9K5_9DEIO	Q4h9k5 deinococcus
218	24	85.7	342	1	PLSX_LEGPA	Q5xsh3 legionella
219	24	85.7	342	1	PLSX_LEGPH	Q5zvp8 legionella
220	24	85.7	342	1	PLSX_LEGPL	Q5wv6 legionella
221	24	85.7	344	1	SYFA_OCEIH	Q8eph4 oceanobacil
222	24	85.7	345	2	Q5WLF7_BACSK	Q5wlf7 bacillus cl
223	24	85.7	347	2	Q4N2G9_THEPA	Q4n2g9 theileria p
224	24	85.7	347	2	Q4U9T3_THEAN	Q4u9t3 theileria a
225	24	85.7	351	2	Q23287_CAEEL	Q23287 caenorhabdi
226	24	85.7	351	2	Q7UNN0_RHOBA	Q7unn0 rhodopirell
227	24	85.7	358	2	Q6C137_YARLI	Q6ci37 yarrowia li
228	24	85.7	360	1	RUVB_COREF	Q8fcp5 corynebacte
229	24	85.7	360	1	RUVB_CORGL	Q9ae09 corynebacte
230	24	85.7	363	1	KCO1_ARATH	Q8lbl1 arabidopsis
231	24	85.7	369	2	Q49641_ARATH	Q49641 arabidopsis
232	24	85.7	373	2	Q5Z200_NOCAE	Q5z200 nocardia fa
233	24	85.7	374	2	Q6FUT0_CANGA	Q6fut0 candida gla
234	24	85.7	375	2	Q9XGI2_LYCES	Q9xgi2 lycopersico
235	24	85.7	376	1	TRUC_ERWCA	Q47417 erwinia car
236	24	85.7	378	2	Q6YU88_ORYSA	Q6yus8 oryza sativ
237	24	85.7	381	2	Q5V294_HALMA	Q5v294 haloarcula
238	24	85.7	383	2	Q7U9U2_SYNPM	Q7u9u2 synchococc
239	24	85.7	384	2	Q88T33_LACPL	Q88t33 lactobacill
240	24	85.7	385	2	Q8LFA6_ARATH	Q8lfa6 arabidopsis
241	24	85.7	385	2	Q93YR5_ARATH	Q93yr5 arabidopsis
242	24	85.7	388	2	Q6NGG9_CORDI	Q6ngg9 corynebacte
243	24	85.7	403	1	Y4635_PSEPK	Q88e33 pseudomonas
244	24	85.7	405	2	Q6CTQ1_KLULA	Q6ctq1 kluyveromyc
245	24	85.7	407	2	Q6A6I2_PROAC	Q6a6i2 propionibac
246	24	85.7	416	2	Q8ELE2_OCEIH	Q8ele2 oceanobacil
247	24	85.7	423	2	Q8XQB6_RALSO	Q8xqb6 ralstonia s
248	24	85.7	429	2	Q5A6B5_CANAL	Q5a6b5 candida alb
249	24	85.7	431	2	Q6AM80_DESPS	Q6am80 desulfotale
250	24	85.7	432	2	Q5N2A4_SYNPM	Q5n2a4 synchococc

251	24	85.7	439	2	Q8EDJ1_SHEON	Q8edj1 shewanella	324	24	85.7	650	2	Q4RF57_TETNG	Q4rf57 tetraodon n
252	24	85.7	446	2	Q61TK1_CAEBR	Q61tk1 caenorhabdi	325	24	85.7	651	2	Q7U3Y4_SYNXP	Q7u3y4 synechococc
253	24	85.7	448	2	Q7PMI2_ANOGA	Q7pmi2 anopheles g	326	24	85.7	654	1	HSP77_YEAST	Pi2398 saccharomyc
254	24	85.7	448	2	Q9GZH2_CAEL	Q9gzh2 caenorhabdi	327	24	85.7	654	2	Q8Y0H3_RALSO	Q8y0h3 ralstonia s
255	24	85.7	452	2	Q5BGK3_EMENI	Q5bgk3 aspergillus	328	24	85.7	659	2	Q554W2_DICDI	Q554w2 dictyosteli
256	24	85.7	453	1	TOLB_GLUOX	Q5fqbp gluconobact	329	24	85.7	671	2	Q97209_LEIMA	Q97209 leishmania
257	24	85.7	459	2	Q32384_SYN7	Q32384 synechococc	330	24	85.7	688	2	Q5W7N9_BACTU	Q5w7n9 bacillus th
258	24	85.7	459	2	Q5N3A0_SYN6	Q5n3a0 synechococc	331	24	85.7	693	1	EFGL_BORGA	Q660Y4 borrelia ga
259	24	85.7	462	1	SAHH_ROSDE	Q9zna5 roseobacter	332	24	85.7	699	2	Q6NZX0_BRARE	Q6nzx0 brachydanio
260	24	85.7	465	1	C85A1_ARATH	Q9fma5 arabidopsis	333	24	85.7	701	2	Q74DM6_GEOSL	Q74dm6 geobacter s
261	24	85.7	465	1	C85A2_ARATH	Q940v4 arabidopsis	334	24	85.7	710	1	TRIM9_HUMAN	Q9c026 homo sapien
262	24	85.7	466	2	Q7PNH9_ANOGA	Q7pnh9 anopheles g	335	24	85.7	710	1	TRIM9_RAT	Q912y8 rattus norv
263	24	85.7	467	2	Q6CT29_KLULA	Q6ctt29 kluyveromyc	336	24	85.7	713	2	Q4REG0_TETNG	Q4reg0 tetraodon n
264	24	85.7	469	2	Q06920_YEAST	Q06920 saccharomyc	337	24	85.7	724	2	Q4R0X3_STRRI	Q4r0x3 streptomyce
265	24	85.7	469	2	Q4WKH9_ASPFU	Q4wkh9 aspergillus	338	24	85.7	731	2	Q7SDQ6_NEUCR	Q7sdq6 neurospora
266	24	85.7	469	2	Q5DCR4_SCHJA	Q5dcr4 schistosoma	339	24	85.7	733	2	Q7UW15_RHOBA	Q7uw15 rhodopirell
267	24	85.7	470	2	Q5JST2_HUMAN	Q5jst2 homo sapien	340	24	85.7	745	2	Q4KF75_PSEF5	Q4kf75 pseudomonas
268	24	85.7	471	1	RUVB2_YEAST	Q12464 saccharomyc	341	24	85.7	750	2	Q8LJZ2_SORBI	Q8ljz2 sorghum bic
269	24	85.7	476	2	Q6FSP1_CANGA	Q6fsf1 candida gla	342	24	85.7	754	2	Q6G162_BARQU	Q6g162 bartonella
270	24	85.7	476	2	Q8BKT9_MOUSE	Q8bkt9 mus musculu	343	24	85.7	754	2	Q6G3Z6_BARHE	Q6g3z6 bartonella
271	24	85.7	480	2	Q6BSB8_DEBHA	Q6bsb8 debaryomyce	344	24	85.7	762	2	Q7S9Z5_NEUCR	Q7s9z5 neurospora
272	24	85.7	480	2	Q5XVA8_ARATH	Q5xva8 arabidopsis	345	24	85.7	781	2	Q8U9K3_AGR75	Q8u9k3 agrobacteri
273	24	85.7	481	2	Q9V3K3_DROME	Q9v3k3 drosophila	346	24	85.7	788	2	Q7CT60_AGR75	Q7ct60 agrobacteri
274	24	85.7	482	2	Q8WSG2_9EUKA	Q8ws92 acraasis ros	347	24	85.7	791	2	Q73U76_MYCPA	Q73u76 mycobacteri
275	24	85.7	498	2	Q5AGZ9_CANAL	Q5agz9 candida alb	348	24	85.7	799	2	Q88F20_PSEPK	Q88f20 pseudomonas
276	24	85.7	501	2	Q41FW5_GIBZE	Q41fw5 gibberella	349	24	85.7	806	2	Q6BRK5_DEBHA	Q6brk5 debaryomyce
277	24	85.7	505	1	BRE2_YEAST	P43132 saccharomyc	350	24	85.7	813	2	Q9FJAJ_ARATH	Q9fjai arabidopsis
278	24	85.7	506	2	Q8EH17_SHEON	Q8eh17 shewanella	351	24	85.7	817	1	TRIM9_MOUSE	Q8c7m3 mus musculu
279	24	85.7	511	2	Q5BC23_EMENI	Q5bcz3 aspergillus	352	24	85.7	831	2	Q5B3M2_EMENI	Q5b3m2 aspergillus
280	24	85.7	512	1	NIR_SYN7	P39661 synechococc	353	24	85.7	842	2	Q82MV5_STRAW	Q82mv5 streptomyce
281	24	85.7	512	2	Q5NSB8_SYN6	Q5nsb8 synechococc	354	24	85.7	843	2	Q4NMV7_9DELT	Q4nmv7 anaeromycob
282	24	85.7	513	2	Q5TBD6_HUMAN	Q5tbd6 homo sapien	355	24	85.7	843	2	Q72RN5_LEPIC	Q72rn5 leptospira
283	24	85.7	517	2	Q9FIH2_ARATH	Q9fih2 arabidopsis	356	24	85.7	884	2	Q5GU97_XANOR	Q5gu97 xanthomonas
284	24	85.7	521	2	Q4UD07_THEAN	Q4ud07 theileria a	357	24	85.7	924	2	Q59QN2_CANAL	Q59qn2 candida alb
285	24	85.7	527	2	Q98C54_RHILO	Q98c54 rhizobium l	358	24	85.7	936	2	Q4RJZ7_TETNG	Q4rjz7 tetraodon n
286	24	85.7	530	2	Q9DW52_RCMVM	Q9dw52 rat cytomeg	359	24	85.7	952	2	Q7XN31_ORYSA	Q7xn31 oryza sativ
287	24	85.7	535	2	Q7Q4U7_ANOGA	Q7q4u7 anopheles g	360	24	85.7	960	2	Q7N7N1_PHOLL	Q7n7n1 photorhabdu
288	24	85.7	538	2	Q95K16_MACFA	Q95k16 macaca fasc	361	24	85.7	966	2	Q4NMD6_9MYCC	Q4nmd6 arthrobacte
289	24	85.7	541	2	Q983Q8_RHILO	Q983q8 rhizobium l	362	24	85.7	988	2	Q600R9_MYCHY	Q600r9 mycoplasma
290	24	85.7	543	2	Q8E5B1_STR43	Q8e5b1 streptococc	363	24	85.7	993	2	Q9U4F8_DROME	Q9u4f8 drosophila
291	24	85.7	551	1	KIF12_HUMAN	Q96fn5 homo sapien	364	24	85.7	997	2	Q9GL4_CAEEL	Q9gyl4 caenorhabdi
292	24	85.7	560	2	Q7UWH7_RHOBA	Q7uwh7 rhodopirell	365	24	85.7	1000	1	Y041_SYNY3	Q55445 synechocyst
293	24	85.7	564	2	Q7S0Z2_NEUCR	Q7s0z2 neurospora	366	24	85.7	1000	2	Q54A85_SYNY3	Q54a85 synechocyst
294	24	85.7	564	2	Q4KCV3_PSEF5	Q4kcv3 pseudomonas	367	24	85.7	1027	2	Q695H0_CHLRE	Q695h0 chlamydomon
295	24	85.7	568	2	Q9ISN0_PSEAB	Q9isn0 pseudomonas	368	24	85.7	1061	2	Q4XQ34_PLACH	Q4xq34 plasmodium
296	24	85.7	569	2	Q98KA7_RHILO	Q98ka7 rhizobium l	369	24	85.7	1061	2	Q8RNY3_HAFAL	Q8rny3 hafnia alve
297	24	85.7	573	2	Q4SFK8_TETNG	Q4sfk8 tetraodon n	370	24	85.7	1070	2	Q6I5X9_ORYSA	Q6i5x9 oryza sativ
298	24	85.7	585	2	Q9V3M9_DROME	Q9v3m9 drosophila	371	24	85.7	1087	2	Q4HWP8_GIBZE	Q4hwp8 gibberella
299	24	85.7	599	2	Q7NYB6_CHRVO	Q7nyb6 chromobacte	372	24	85.7	1103	2	Q7V6Q7_PROMM	Q7v6q7 prochloroco
300	24	85.7	600	2	Q9HSM3_HALSA	Q9hsm3 halobacteri	373	24	85.7	1104	2	Q9XEQ2_SORBI	Q9xeq2 sorghum bic
301	24	85.7	600	2	Q4PE22_USTMA	Q4pe22 ustilago ma	374	24	85.7	1105	2	Q4HWT9_GIBZE	Q4hwt9 gibberella
302	24	85.7	600	2	Q869T6_DICDI	Q869t6 dictyosteli	375	24	85.7	1140	2	Q4WNA6_ASPFU	Q4wna6 aspergillus
303	24	85.7	600	2	Q5HTL6_CAMJR	Q5htl6 campylobact	376	24	85.7	1175	2	Q9ZPF3_ARATH	Q9zpf3 arabidopsis
304	24	85.7	600	2	Q9PN51_CAMJE	Q9pn51 campylobact	377	24	85.7	1175	2	Q9SMU3_ARATH	Q9smu3 arabidopsis
305	24	85.7	600	2	Q8EV40_MYCPE	Q8ev40 mycoplasma	378	24	85.7	1179	2	Q7SAB3_NEUCR	Q7sab3 neurospora
306	24	85.7	602	2	Q8F426_LEPIN	Q8f426 leptospira	379	24	85.7	1195	2	Q4ICQ6_GIBZE	Q4icq6 gibberella
307	24	85.7	603	2	Q9L217_STRCO	Q9l217 streptomyce	380	24	85.7	1195	2	Q4ICQ6_GIBZE	Q9sn55 arabidopsis
308	24	85.7	604	2	Q55F14_DICDI	Q55f14 dictyosteli	381	24	85.7	1269	2	Q7UH58_RHOBA	Q7uh58 rhodopirell
309	24	85.7	606	2	Q4TR89_9SPHN	Q4tr89 erythrobact	382	24	85.7	1326	1	CI093_HUMAN	Q6tfl3 homo sapien
310	24	85.7	610	2	Q6GLX6_XENLA	Q6glx6 xenopus lae	383	24	85.7	1358	2	Q6L3H0_SOLDE	Q6l3h0 solanum dem
311	24	85.7	615	2	Q512R9_ENTHI	Q512r9 entamoeba h	384	24	85.7	1484	2	Q5A6J4_CANAL	Q5a6j4 candida alb
312	24	85.7	621	2	Q7YXJ6_TRYCR	Q7yxj6 trypanosoma	385	24	85.7	1489	2	Q23302_ARATH	Q23302 arabidopsis
313	24	85.7	624	2	Q8G5P3_BIFLO	Q8g5p3 bifidobacte	386	24	85.7	1545	2	Q5KKG1_CRYNE	Q5kk91 cryptococcu
314	24	85.7	628	1	FTSH3_SYNY3	P73437 synechocyst	387	24	85.7	1545	2	Q55VSB_CRYNE	Q55v88 cryptococcu
315	24	85.7	628	2	Q66DT0_YERPS	Q66dt0 yersinia ps	388	24	85.7	1633	2	Q81617_ARATH	Q81617 arabidopsis
316	24	85.7	628	2	Q8ZC69_YERPE	Q8zcc69 yersinia pe	389	24	85.7	1702	2	Q7PS39_ANOGA	Q7ps39 anopheles g
317	24	85.7	634	2	Q95JX7_MACFA	Q95jx7 macaca fasc	390	24	85.7	1711	2	Q59DT7_DROME	Q59dt7 drosophila
318	24	85.7	640	2	Q5TKI6_ORYSA	Q5tki6 oryza sativ	391	24	85.7	1756	2	Q6AWK8_DROME	Q6awk8 drosophila
319	24	85.7	646	2	Q5TBE0_HUMAN	Q5tbe0 homo sapien	392	24	85.7	1850	2	Q59DT8_DROME	Q59dt8 drosophila
320	24	85.7	646	2	Q8JH68_BRARE	Q8jh68 brachydanio	393	24	85.7	1851	1	LAP4_DROME	Q7kry7 drosophila
321	24	85.7	648	2	Q9HUH9_PSEAE	Q9huh9 pseudomonas	394	24	85.7	1986	2	Q7RFL5_PLAYO	Q7rfl5 plasmodium
322	24	85.7	650	2	Q4KD73_PSEF5	Q4kd73 pseudomonas	395	24	85.7	2023	2	Q91632_9VIRU	Q91632 cherry gree
323	24	85.7	650	2	Q88IY0_PSEPK	Q88iy0 pseudomonas	396	24	85.7	2937	2	Q7N7K0_PHOLL	Q7n7k0 photorhabdu

397	24	85.7	2961	2	Q7MB60_PHOLL	Q7mb60 photorhabdu	470	23	82.1	204	2	Q4T3U3_TETNG	Q4t3u3 tetraodon n
398	24	85.7	3027	2	Q7MB03_PHOLL	Q7mb03 photorhabdu	471	23	82.1	205	2	Q5V7J6_HALMA	Q5v7j6 haloarcula
399	24	85.7	3056	1	POLG_BYMV	p17765 b genome po	472	23	82.1	205	2	Q889C8_PSESM	Q889c8 pseudomonas
400	24	85.7	3056	2	Q65852_BYMV	Q65852 bean yellow	473	23	82.1	209	2	Q5NFZ1_FRATT	Q5nfz1 francisella
401	24	85.7	3056	2	Q65892_BYMV	Q65892 bean yellow	474	23	82.1	211	2	Q8X147_ALTPR	Q8x147 alternaria
402	24	85.7	3072	2	Q92645_CYVV	Q92645 clover yell	475	23	82.1	211	2	Q98EE8_RHILO	Q98ee8 rhizobium l
403	24	85.7	3501	2	Q8Y106_RALSO	Q8y106 ralstonia s	476	23	82.1	212	2	Q5WZS5_LEGPL	Q5wzs5 legionella
404	24	85.7	3552	2	Q8XSD6_RALSO	Q8xsd6 ralstonia s	477	23	82.1	212	2	Q5X8D0_LEGPA	Q5x8d0 legionella
405	24	85.7	4582	2	Q7N911_PHOLL	Q7n911 photorhabdu	478	23	82.1	212	2	Q8FLT7_COREF	Q8flt7 corynebacte
406	24	85.7	4602	2	Q6V7L9_9CAUD	Q6v7l9 burkholderi	479	23	82.1	213	2	Q8DM96_SYNEL	Q8dm96 synechococc
407	24	85.7	7538	2	Q7UNV3_RHOBA	Q7unv3 rhodopirell	480	23	82.1	216	1	MTTC_METBA	Q93659 methanosarc
408	23	82.1	60	2	Q9TX85_CAEEL	Q9tx85 caenorhabdi	481	23	82.1	216	2	Q6FFI8_ACIAD	Q6ffi8 acinetobact
409	23	82.1	61	2	Q8J152_9PLEO	Q8j152 leptosphaer	482	23	82.1	218	1	MTTC2_METMA	P58983 methanosarc
410	23	82.1	83	1	GVPG2_HALSA	Q8j152 leptosphaer	482	23	82.1	218	1	MTTC2_METMA	Q5wtd8 legionella
411	23	82.1	89	2	Q4RCE2_TETNG	Q4rce2 tetraodon n	483	23	82.1	220	2	Q5WVJ8_LEGPL	Q5wvj8 legionella
412	23	82.1	91	1	HS30E_XENLA	P30220 xenopus lae	484	23	82.1	220	2	Q5X1M9_LEGPA	Q5x1m9 legionella
413	23	82.1	93	2	Q88EF2_PSEPK	Q88ef2 pseudomonas	485	23	82.1	220	2	Q5X462_LEGPA	Q5x462 legionella
414	23	82.1	95	2	Q4FLQ6_9RICK	Q4flq6 candidatus	486	23	82.1	220	2	Q5ZS59_LEGPH	Q5z559 legionella
415	23	82.1	97	2	Q65LJ0_BACLD	Q65lj0 bacillus li	487	23	82.1	220	2	Q5ZUE7_LEGPH	Q5zue7 legionella
416	23	82.1	102	2	Q71IF5_LACDL	Q71if5 lactobacill	488	23	82.1	221	2	Q919T2_9INFA	Q919t2 influenza a
417	23	82.1	102	2	Q60B40_METCA	Q60b40 methylococc	489	23	82.1	221	2	Q919T3_9INFA	Q919t3 influenza a
418	23	82.1	110	2	Q60C52_METCA	Q60c52 methylococc	490	23	82.1	221	2	Q7NCM0_GLOVI	Q7ncm0 gloeobacter
419	23	82.1	115	2	Q8EIS7_SHEON	Q8eis7 shewanella	491	23	82.1	226	2	Q5YW85_NOCFA	Q5yw85 nocardia fa
420	23	82.1	119	2	Q5UZV1_HALMA	Q5uzv1 haloarcula	492	23	82.1	228	2	Q81PL3_BACAN	Q81pl3 bacillus an
421	23	82.1	121	2	Q6FCV2_ACIAD	Q6fcv2 acinetobact	493	23	82.1	228	2	Q6HHU8_BACHK	Q6hhu8 bacillus th
422	23	82.1	122	1	RL18_RHOBA	Q7un04 rhodopirell	494	23	82.1	228	2	Q7UH08_RHOBA	Q7uh08 rhodopirell
423	23	82.1	129	2	Q825V1_STRAW	Q825v1 streptomyce	495	23	82.1	233	2	Q6UE31_9LACO	Q6ue31 lactobacill
424	23	82.1	132	2	Q9AN59_BRAJA	Q9an59 bradyrhizob	496	23	82.1	236	2	Q74886_SCHPO	O74886 schizosacch
425	23	82.1	135	2	Q8TPW1_METAC	Q8tpw1 methanosarc	497	23	82.1	237	2	Q4QAP9_LEIMA	Q4qap9 leishmania
426	23	82.1	140	2	Q9SAL1_ARATH	Q9sal1 arabidopsis	498	23	82.1	237	2	Q4QAP9_LEIMA	Q4qap9 leishmania
427	23	82.1	143	2	Q87FW9_VIBPA	Q87fw9 vibrio para	499	23	82.1	238	2	O81922_CAPAN	O81922 capsicum an
428	23	82.1	143	2	Q6SUZ8_9BACU	Q6suz8 achaea jana	500	23	82.1	241	2	Q8RM80_LEGPN	Q8rm80 legionella
429	23	82.1	144	2	Q6K3C6_ORYSA	Q6k3c6 oryza sativ	501	23	82.1	242	2	Q604M9_METCA	Q604m9 methylococc
430	23	82.1	147	2	Q825J2_STRAW	Q825j2 streptomyce	502	23	82.1	244	2	Q87EE8_XYLFT	Q87ee8 xylella fas
431	23	82.1	148	2	Q8PJ88_XANAC	Q8pj88 xanthomonas	503	23	82.1	245	2	Q4NUG0_9DELT	Q4nug0 anaeromycob
432	23	82.1	149	2	Q8ZH71_YERPE	Q8zh71 yersinia pe	504	23	82.1	248	2	Q05077_STRCL	Q05077 streptomyce
433	23	82.1	149	2	Q667I4_YERPS	Q667i4 yersinia ps	505	23	82.1	249	1	UBIG_RHILO	Q32722 paenibacill
434	23	82.1	150	2	Q50S28_ENTHI	Q50s28 entamoeba h	506	23	82.1	251	2	O32722_PAEPO	Q32722 paenibacill
435	23	82.1	150	2	Q66086_9ALPH	Q66086 canine herp	507	23	82.1	262	2	Q5Z8E2_ORYSA	Q5z8e2 oryza sativ
436	23	82.1	151	2	Q87GT1_VIBPA	Q87gt1 vibrio para	508	23	82.1	263	2	Q8PPV1_XANAC	Q8ppv1 xanthomonas
437	23	82.1	153	2	Q4H6G1_9DEIO	Q4hg61 deinococcus	509	23	82.1	264	2	Q66V92_9ACTO	Q66v92 nonomuraea
438	23	82.1	154	1	PFDA_AERPE	Q9ydz28 aeropyrum p	510	23	82.1	265	2	Q5WKM2_BACSK	Q5wkm2 bacillus cl
439	23	82.1	158	2	Q5YS89_NOCFA	Q5ys89 nocardia fa	511	23	82.1	265	2	O07412_MYCTU	Q7u2p1 mycobacteri
440	23	82.1	161	2	Q6I8P6_XENLA	Q6i8p6 xenopus lae	512	23	82.1	265	2	O07412_MYCTU	O07412 mycobacteri
441	23	82.1	167	2	Q8VLN8_9PSED	Q8vln8 pseudomonas	513	23	82.1	268	2	Q9AEZ0_9ACTO	Q9aez0 frankia sp.
442	23	82.1	167	2	Q6G502_BARHE	Q6g502 bartonella	514	23	82.1	269	2	Q5V0B1_HALMA	Q5v0b1 haloarcula
443	23	82.1	167	2	Q8ELM3_OCEIH	Q8elm3 oceanobacil	515	23	82.1	269	2	Q4SKD7_TETNG	Q4skd7 tetraodon n
444	23	82.1	167	2	Q4LZK2_9BURK	Q4lzk2 burkholderi	516	23	82.1	275	2	Q9HDD4_9PLEO	Q9hdd4 alternaria
445	23	82.1	168	2	Q7Q9Y1_ANOGA	Q7q9y1 anopheles g	517	23	82.1	275	2	Q9HES6_9PLEO	Q9hes6 alternaria
446	23	82.1	169	2	Q50NW5_ENTHI	Q50nw5 entamoeba h	518	23	82.1	275	2	Q7UZC5_RHOBA	Q7uzc5 rhodopirell
447	23	82.1	170	2	Q5UQC6_MIMIV	Q5uqc6 mimivirus	519	23	82.1	278	2	Q8X146_9PLEO	Q8x146 alternaria
448	23	82.1	171	2	Q9Z7K5_CHLPN	Q9z7k5 chlamydia p	520	23	82.1	278	2	Q8BCC5_9INFA	Q8bcc5 influenza a
449	23	82.1	172	2	Q5L767_CHLAB	Q5l767 chlamydophi	521	23	82.1	278	2	Q8BCC6_9INFA	Q8bcc6 influenza a
450	23	82.1	172	2	Q824U7_CHLCV	Q824u7 chlamydophi	522	23	82.1	278	2	Q8BCC7_9INFA	Q8bcc7 influenza a
451	23	82.1	173	2	O84683_CHLTR	O84683 chlamydia t	523	23	82.1	278	2	Q8BCC8_9INFA	Q8bcc8 influenza a
452	23	82.1	174	2	Q7V672_PROMM	Q7v672 prochloroco	524	23	82.1	278	2	Q8BCC9_9INFA	Q8bcc9 influenza a
453	23	82.1	177	2	Q9PLP8_CHLMU	Q9plp8 chlamydia m	525	23	82.1	278	2	Q8BCD0_9INFA	Q8bcd0 influenza a
454	23	82.1	180	2	Q9PAL1_XYLFA	Q9pal1 xylella fas	526	23	82.1	278	2	Q8BCD1_9INFA	Q8bcd1 influenza a
455	23	82.1	180	2	Q67JN3_SYMTH	Q67jn3 symbiobacte	527	23	82.1	278	2	Q8BCD2_9INFA	Q8bcd2 influenza a
456	23	82.1	180	2	Q877M8_XYLFT	Q877m8 xylella fas	528	23	82.1	278	2	Q8BCD3_9INFA	Q8bcd3 influenza a
457	23	82.1	181	2	Q8DH88_SYNEL	Q8dh88 synechococc	529	23	82.1	283	2	Q8BCD4_9INFA	Q8bcd4 influenza a
458	23	82.1	184	2	Q65PD8_BACLD	Q65pd8 bacillus li	530	23	82.1	289	2	Q5JW01_HUMAN	Q5jw01 homo sapien
459	23	82.1	190	2	Q72QY6_LEPIC	Q72qy6 leptospira	531	23	82.1	289	2	Q5GMG5_MOUSE	Q5gmgs mus musculu
460	23	82.1	190	2	Q8F4V4_LEPIN	Q8f4v4 leptospira	532	23	82.1	291	2	Q7PSX0_ANOGA	Q7psx0 anopheles g
461	23	82.1	192	1	Y4GC_RHISN	P55459 rhizobium s	533	23	82.1	297	2	Q86EA5_SCHJA	Q86ea5 schistosoma
462	23	82.1	193	2	Q859F3_9VIRU	Q859f3 uncultured	534	23	82.1	297	2	Q88LC6_PSEPK	Q88lc6 pseudomonas
463	23	82.1	193	2	Q8EDR6_SHEON	Q8edr6 shewanella	535	23	82.1	298	2	Q6U737_VIBFI	Q6u737 vibrio fisc
464	23	82.1	198	2	Q859N6_9VIRU	Q859n6 uncultured	536	23	82.1	301	2	O24760_HYPME	O24760 hyphomicrob
465	23	82.1	203	2	Q92C91_LISIN	Q92c91 listeria in	537	23	82.1	302	2	Q6LSS6_PHOPR	Q6lss6 photobacter
466	23	82.1	204	2	Q7PE10_ANOGA	Q7pe10 anopheles g	538	23	82.1	302	2	Q88YC3_LACPL	Q88yc3 lactobacill
467	23	82.1	204	2	Q4ZUD7_PSESY	Q4zud7 pseudomonas	539	23	82.1	303	1	PLSC_YEAST	P33333 saccharomyc
468	23	82.1	204	2	Q8EX26_MYCPE	Q8ex26 mycoplasma	540	23	82.1	303	2	Q6B238_YEAST	Q6b238 saccharomyc
469	23	82.1	204	2	Q6D313_ERWCT	Q6d313 erwinia car	541	23	82.1	303	2	Q97I47_CLOAB	Q97i47 clostridium
							542	23	82.1	306	2	Q9W5G7_DROME	Q9w5g7 drosophila

689	23	82.1	409	2	Q80791_ARATH	Q80791_arabidopsis	762	23	82.1	457	2	Q4NT81_9DELT	Q4nt81_anaeromyxob
690	23	82.1	409	2	Q6YZ60_ORYSA	Q6yz60_oryza sativ	763	23	82.1	457	2	Q6DDN1_XENLA	Q6ddn1_xenopus lae
691	23	82.1	409	2	Q8GG5_GEOXA	Q8gg5_geobacillus	764	23	82.1	459	2	Q74L28_LACJO	Q74l28_lactobacillio
692	23	82.1	411	2	Q8QVC7_9REOV	Q8qvc7_eyach virus	765	23	82.1	459	2	Q58EL7_BRARE	Q58el7_brachydanio
693	23	82.1	411	2	Q5RH89_BRARE	Q5rh89_brachydanio	766	23	82.1	460	1	RAB31I_RAT	Q62739_rattus norv
694	23	82.1	413	2	Q9VJL7_DROME	Q9vj17_drosophila	767	23	82.1	460	1	SYC_PSESM	Q87yq2_pseudomonas
695	23	82.1	414	2	Q9NK60_DROME	Q9nk60_drosophilala	768	23	82.1	460	2	Q5R5E6_PONPY	Q5r5e6_pongo pygma
696	23	82.1	415	2	Q4J4F0_AZOVI	Q4j4f0_azotobacter	769	23	82.1	460	2	Q4K9S1_PSEF5	Q4k9s1_pseudomonas
697	23	82.1	415	2	Q9JIE4_MOUSE	Q9jie4_mus musculus	770	23	82.1	460	2	Q8E9P4_SHEON	Q8e9p4_shewanella
698	23	82.1	418	2	Q72TP7_LEPIC	Q72tp7_leptospira	771	23	82.1	461	2	Q82ZY2_ENTFA	Q82zy2_enterococcu
699	23	82.1	418	2	Q8F1J2_LEPIN	Q8f1j2_leptospira	772	23	82.1	464	2	Q4I948_GIBZE	Q4i948_gibberella
700	23	82.1	420	2	Q4JXS6_CORJK	Q4jxs6_corynebacte	773	23	82.1	464	2	Q2I1794_CAEEL	Q2i1794_caenorhabdi
701	23	82.1	423	2	Q75TW9_BACHD	Q75tw9_bacillus ha	774	23	82.1	465	2	Q524T0_MAGGR	Q524t0_magnaporthe
702	23	82.1	424	2	Q4ZSW7_PSESY	Q4zsw7_pseudomonas	775	23	82.1	465	2	Q4MZR9_THEPA	Q4mzr9_theileria p
703	23	82.1	424	2	Q5L579_CHLAB	Q5l579_chlamydophi	776	23	82.1	465	2	Q7V3A5_PROMP	Q7v3a5_prochloroco
704	23	82.1	425	2	Q7WX91_ALCEU	Q7wx91_alcaligenes	777	23	82.1	465	2	Q6P8E3_XENTR	Q6p8e3_xenopus tro
705	23	82.1	425	2	Q4M0Q4_9BURK	Q4m0q4_burkholderi	778	23	82.1	469	2	Q9L3C1_AMYMD	Q9l3c1_amycolatops
706	23	82.1	425	2	Q5QVS9_IDILO	Q5qvs9_idiomarina	780	23	82.1	470	1	LEU2_BACSK	Q5wen5_bacillus cl
707	23	82.1	425	2	Q5O3R3_BRARE	Q5o3r3_brachydanio	781	23	82.1	470	2	Q9GZI6_CAEEL	Q9gzi6_caenorhabdi
708	23	82.1	426	2	Q4UK60_RICFE	Q4uk60_rickettsia	781	23	82.1	470	2	Q7U1F9_MYCBO	Q7u1f9_mycobacteri
709	23	82.1	426	2	Q68VY5_RICTY	Q68vy5_rickettsia	783	23	82.1	470	2	P95042_MYCTU	P95042_mycobacteri
710	23	82.1	426	2	Q9A7P2_CAUCR	Q9a7p2_caulobacter	783	23	82.1	471	1	FBX3_HUMAN	Q9uk99_homo sapien
711	23	82.1	426	2	Q9ZCH8_RICPR	Q9zch8_rickettsia	784	23	82.1	471	2	Q9H0V2_HUMAN	Q9h0v2_homo sapien
712	23	82.1	427	2	Q7Q6B4_ANOGA	Q7q6b4_anopheles g	785	23	82.1	471	2	Q86X90_HUMAN	Q86x90_homo sapien
713	23	82.1	427	2	Q75TN6_BACFI	Q75tn6_bacillus fi	786	23	82.1	471	2	Q8FE74_ECOL6	Q8fe74_escherichia
714	23	82.1	427	2	Q75TQ1_BACAO	Q75tq1_bacillus al	787	23	82.1	473	2	Q9RQE6_9MICO	Q9rqe6_xylanimicro
715	23	82.1	427	2	Q75TX1_BACHD	Q75tx1_bacillus ha	788	23	82.1	474	2	Q70LR1_LEGPN	Q70lr1_legionella
716	23	82.1	427	2	Q75TX3_BACHD	Q75tx3_bacillus ha	789	23	82.1	474	2	Q5X4V5_LEGPA	Q5x4v5_legionella
717	23	82.1	427	2	Q75TX6_BACHD	Q75tx6_bacillus ha	790	23	82.1	475	2	Q4FQP5_9GAMM	Q4fqp5_psychrobact
718	23	82.1	427	2	Q75TX7_BACHD	Q75tx7_bacillus ha	791	23	82.1	476	1	RAB31I_HUMAN	Q96qf0_homo sapien
719	23	82.1	427	2	Q75TZ5_BACHD	Q75tz5_bacillus ha	792	23	82.1	476	2	Q6BUYI_DEBHA	Q6buy1_debaryomyce
720	23	82.1	427	2	Q75TZ7_BACHD	Q75tz7_bacillus ha	793	23	82.1	476	2	Q5RFP1_PONPY	Q5rfp1_pongo pygma
721	23	82.1	427	2	Q75TZ8_BACHD	Q75tz8_bacillus ha	794	23	82.1	480	1	FBX3_MOUSE	Q9dc63_mus musculu
722	23	82.1	427	2	Q73V81_MYCPA	Q73v81_mycobacteri	795	23	82.1	480	2	Q5UZE8_HALMA	Q5uze8_haloarcula
723	23	82.1	427	2	Q89K16_BRAJA	Q89k16_bradyrhizob	796	23	82.1	480	2	Q8C7I0_MOUSE	Q8c7i0_mus musculu
724	23	82.1	427	2	Q9K5P7_BACHD	Q9k5p7_bacillus ha	797	23	82.1	480	2	Q9D2J5_MOUSE	Q9d2j5_mus musculu
725	23	82.1	428	1	R SMB_SHEON	Q8ekr0_shewanella	798	23	82.1	480	2	Q8CEJ0_MOUSE	Q8cej0_mus musculu
726	23	82.1	428	2	Q7P9T3_RICSI	Q7p9t3_rickettsia	799	23	82.1	481	2	Q873C7_NEUCR	Q873c7_neurospora
727	23	82.1	428	2	Q92GD7_RICCN	Q92gd7_rickettsia	800	23	82.1	483	2	Q4ZQS6_PSESY	Q4zqs6_pseudomonas
728	23	82.1	429	2	Q5ZKG9_CHICK	Q5zkg9_gallus gall	801	23	82.1	484	2	Q7ZON8_PARTE	Q7z0n8_paramecium
729	23	82.1	429	2	Q7ZWA9_BRARE	Q7zwa9_brachydanio	802	23	82.1	485	2	Q89CL4_BRAJA	Q89cl4_bradyrhizob
730	23	82.1	430	2	Q5RBZ1_PONPY	Q5rbz1_pongo pygma	803	23	82.1	492	2	Q4UBZ8_THEAN	Q4ubz8_theileria a
731	23	82.1	431	2	Q66II6_XENTR	Q66ii6_xenopus tro	804	23	82.1	492	2	Q76EA6_PSESG	Q76ea6_pseudomonas
732	23	82.1	431	2	Q7SYG0_ACIBE	Q7syg0_acipenser b	805	23	82.1	492	2	Q6I807_PSESZ	Q6i807_pseudomonas
733	23	82.1	433	2	Q4UHL1_THEAN	Q4uhl1_theileria a	806	23	82.1	492	2	Q884Y5_PSESM	Q884y5_pseudomonas
734	23	82.1	433	2	Q64807_ARATH	Q64807_arabidopsis	807	23	82.1	495	2	Q9HFJ1_NEUCR	Q9h fj1_neurospora
735	23	82.1	434	2	Q832B6_ENTFA	Q832b6_enterococcu	808	23	82.1	501	2	Q5DF74_SCHJA	Q5df74_schistosoma
736	23	82.1	434	2	Q9DRL0_9INFA	Q9drl0_influenza a	809	23	82.1	501	2	Q6KB68_9POTY	Q6kb68_sweet potat
737	23	82.1	434	2	Q9DRL1_9INFA	Q9drl1_influenza a	810	23	82.1	501	2	Q599X8_9POTY	Q599x8_sweet potat
738	23	82.1	434	2	Q9DRL2_9INFA	Q9drl2_influenza a	811	23	82.1	501	2	Q8B323_9POTY	Q8b323_sweet potat
739	23	82.1	434	2	Q8AVM2_XENLA	Q8avm2_xenopus lae	812	23	82.1	501	2	Q8B325_9POTY	Q8b325_sweet potat
740	23	82.1	435	2	Q5P396_AZOSE	Q5p396_azocarcus sp	813	23	82.1	501	2	Q8B326_9POTY	Q8b326_sweet potat
741	23	82.1	437	2	Q4HX18_GIBZE	Q4hx18_gibberella	814	23	82.1	501	2	Q8B327_9POTY	Q8b327_sweet potat
742	23	82.1	437	2	Q87H22_VIBPA	Q87h22_vibrio para	815	23	82.1	502	2	Q6NGC7_CORDI	Q6ngc7_corynebacte
743	23	82.1	440	2	Q86ZW3_LEPMC	Q86zw3_leptosphaer	816	23	82.1	502	2	Q7MMH5_VIBVY	Q7mmh5_vibrio vuln
744	23	82.1	440	2	Q5JVZ9_HUMAN	Q5jvz9_homo sapien	817	23	82.1	503	1	CP7A1_RAT	P18125_rattus norv
745	23	82.1	444	2	Q5XI46_RAT	Q5xi46_rattus norv	818	23	82.1	503	2	Q8UWR4_9INFA	Q8uwr4_influenza a
746	23	82.1	446	2	Q59GS0_HUMAN	Q59gs0_homo sapien	819	23	82.1	504	1	CP7A1_CRIGR	P46634_cricetulus
747	23	82.1	446	2	Q87D83_XYLFT	Q87d83_xylella fas	820	23	82.1	504	2	Q9HQY0_HALSA	Q9h qy0_halobacteri
748	23	82.1	446	2	Q9PBX6_XYLFA	Q9pbx6_xylella fas	821	23	82.1	504	2	Q7REB1_PLAYO	Q7reb1_plasmodium
749	23	82.1	447	2	Q4KEC1_PSEF5	Q4kecl_pseudomonas	822	23	82.1	505	2	Q73UQ4_MYCPA	Q73uq4_mycobacteri
750	23	82.1	448	2	Q4H843_9DEIO	Q4h843_deinococcus	823	23	82.1	506	2	Q9H192_THEAC	Q9hi92_thermoplasm
751	23	82.1	451	2	Q5JW00_HUMAN	Q5jw00_homo sapien	824	23	82.1	508	2	Q4WLA0_ASPFU	Q4wla0_aspergillus
752	23	82.1	451	2	Q4NA38_9MICC	Q4na38_arthrobacter	825	23	82.1	508	2	Q5FQN2_GLUOX	Q5fgn2_gluconobact
753	23	82.1	451	2	Q9A392_CAUCR	Q9a392_caulobacter	826	23	82.1	508	2	Q8DX32_SYNEL	Q8dk32_synechococc
754	23	82.1	451	2	Q5ILH1_CHICK	Q5ilh1_gallus gall	827	23	82.1	509	2	Q4IDJ1_GIBZE	Q4idj1_gibberella
755	23	82.1	452	2	Q4P062_USTMA	Q4p062_ustilago ma	828	23	82.1	513	2	Q5CXR6_CRYPV	Q5cxr6_cryptospori
756	23	82.1	452	2	Q4P721_USTMA	Q4p721_ustilago ma	829	23	82.1	513	2	Q5CIH4_CRYVHO	Q5cih4_cryptospori
757	23	82.1	452	2	Q8EYJ0_LEPIN	Q8eyj0_leptospira	830	23	82.1	513	2	Q6NGC3_CORDI	Q6ngc3_corynebacte
758	23	82.1	452	2	Q72M16_LEPIC	Q72m16_leptospira	831	23	82.1	513	2	Q7NCS3_GLOVI	Q7ncs3_gloeobacter
759	23	82.1	454	2	Q4ZLS9_PSESY	Q4zls9_pseudomonas	832	23	82.1	514	2	Q5FH47_EHRRG	Q5fh47_ehrlichia r
760	23	82.1	454	2	Q7SYQ3_XENLA	Q7syq3_xenopus lae	833	23	82.1	514	2	Q5HC18_EHRRW	Q5hc18_ehrlichia r
761	23	82.1	455	2	Q87UI9_PSESM	Q87ui9_pseudomonas	834	23	82.1	517	2	Q5RJG8_MOUSE	Q5rjg8_mus musculu

835	23	82.1	519	2	Q5B292	EMENI	Q5b292	aspergillus	908	23	82.1	601	2	Q54IK5	DICDI	Q54ik5	dictyosteli
836	23	82.1	522	2	Q83BQ2	COXBU	Q83bq2	coxiella bu	909	23	82.1	604	2	Q4HH51	CAMCO	Q4hh51	campylobact
837	23	82.1	523	2	Q95U04	DROME	Q95u04	drosophila	910	23	82.1	608	2	Q5E4C5	VIBF1	Q5e4c5	vibrio fisc
838	23	82.1	523	2	Q9VP67	DROME	Q9vp67	drosophila	911	23	82.1	609	2	Q8ZZV0	PYRAE	Q8zzv0	pyrobaculum
839	23	82.1	523	2	Q5HGF6	STAAC	Q5hgf6	staphylococ	912	23	82.1	609	2	Q4THX1	TETNG	Q4thx1	tetraodon n
840	23	82.1	525	2	Q89R24	BRAJA	Q89r24	bradyrhizob	913	23	82.1	610	1	UVRC	ECO57	P0a8g2	escherichia
841	23	82.1	527	2	Q05855	MYCTU	Q05855	mycobacteri	914	23	82.1	610	1	UVRC	ECOL6	P0a8g1	escherichia
842	23	82.1	530	1	LNT RHIL0		Q98bj6	rhizobium l	915	23	82.1	610	1	UVRC	ECOLI	P0a8g0	escherichia
843	23	82.1	531	2	Q9HIE2	THEAC	Q9hie2	thermoplasm	916	23	82.1	610	1	UVRC	SALTI	Q825c3	salmonella
844	23	82.1	533	2	Q4INC4	GIBB2E	Q4inc4	gibberella	917	23	82.1	610	1	UVRC	SALTY	O66041	salmonella
845	23	82.1	534	2	Q94FM4	TOBAC	Q94fm4	nicotiana t	918	23	82.1	610	1	UVRC	YERPE	Q8zf52	yersinia pe
846	23	82.1	535	2	Q51IY3	MAGGR	Q51iy3	magnaportha	919	23	82.1	610	2	Q5SNT7	HUMAN	Q5snt7	homo sapien
847	23	82.1	536	2	Q50ME1	ENTHI	Q50me1	entamoeba h	920	23	82.1	610	2	Q57N56	SALCH	Q57n56	salmonella
848	23	82.1	536	2	Q87BL3	XYLFT	Q87bl3	xylella fas	921	23	82.1	610	2	Q5PI10	SALPA	Q5pi10	salmonella
849	23	82.1	536	2	Q9PAS5	XYLFA	Q9pas5	xylella fas	922	23	82.1	610	2	Q66BN5	YERPS	Q66bn5	yersinia ps
850	23	82.1	537	2	Q94FM5	TOBAC	Q94fm5	nicotiana t	923	23	82.1	610	2	Q6D363	ERWCT	Q6d363	erwinia car
851	23	82.1	539	2	Q5H2E3	XANOR	Q5h2e3	xanthomonas	924	23	82.1	610	2	P73400	SYNY3	P73400	synecocyst
852	23	82.1	543	2	Q5A1C8	CANAL	Q5a1c8	candida alb	925	23	82.1	611	1	UVRC	PASMU	P57882	pasteurella
853	23	82.1	544	2	Q4WMW7	ASPFU	Q4wmw7	aspergillus	926	23	82.1	611	2	O69506	MYCLE	O69506	mycobacteri
854	23	82.1	544	2	Q9CCH3	MYCLE	Q9cch3	mycobacteri	927	23	82.1	611	2	O53666	MYCTU	O53666	mycobacteri
855	23	82.1	545	2	Q6BQZ4	DEBHA	Q6bqz4	debaromyce	928	23	82.1	611	2	Q7U2H0	MYCBO	Q7u2h0	mycobacteri
856	23	82.1	548	2	Q71MJ3	CITSI	Q71mj3	citrus sine	929	23	82.1	613	2	P72613	SYNY3	P72613	synecocyst
857	23	82.1	548	2	Q8RVR2	CITPA	Q8rvr2	citrus para	930	23	82.1	619	2	Q4Q268	LEIMA	Q4q268	leishmania
858	23	82.1	555	2	Q6CIU4	KLULA	Q6ciu4	kluveromyc	931	23	82.1	620	2	Q9HXS8	PSEAE	Q9hxs8	pseudomonas
859	23	82.1	555	2	Q9FYU6	9ROSI	Q9fyu6	citrus junco	932	23	82.1	621	2	Q6YZ61	ORYSA	Q6yz61	oryza sativ
860	23	82.1	557	2	Q4L5X8	STAHJ	Q4l5x8	staphylococ	933	23	82.1	626	1	UVRC	SYNY3	P73580	synecocyst
861	23	82.1	557	2	Q6G9T8	STAAS	Q6g9t8	staphylococ	934	23	82.1	627	1	UVRC	ANASP	Q8yti9	anabaena sp
862	23	82.1	557	2	Q6GHG0	STAAR	Q6ghg0	staphylococ	935	23	82.1	631	2	Q4HMT1	CAMLA	Q4hmt1	campylobact
863	23	82.1	557	2	Q7A115	STAAM	Q7a115	staphylococ	936	23	82.1	632	1	STC CLOBE		P26047	clostridium
864	23	82.1	557	2	Q7A5X6	STAAN	Q7a5x6	staphylococ	937	23	82.1	632	2	Q7Z5G9	HUMAN	Q7z5g9	homo sapien
865	23	82.1	557	2	Q99UJ7	STAAM	Q99uj7	staphylococ	938	23	82.1	632	2	Q89KU2	BRADJA	Q89ku2	bradyrhizob
866	23	82.1	559	2	Q51LU9	MAGGR	Q51lu9	magnaportha	939	23	82.1	633	2	Q5NYD2	AZOSE	Q5nyd2	azotarcus sp
867	23	82.1	559	2	Q8H144	ARATH	Q8h144	arabidopsis	940	23	82.1	634	2	Q8H150	ARATH	Q8h150	arabidopsis
868	23	82.1	562	2	Q9Y8D7	9PLEO	Q9y8d7	cochliobolu	941	23	82.1	634	2	Q9FN91	ARATH	Q9fn91	arabidopsis
869	23	82.1	562	2	Q7JK62	CABEL	Q7jk62	caenorhabdi	942	23	82.1	634	2	Q6APJ9	DESPP	Q6apj9	desulfotale
870	23	82.1	565	2	Q6FT63	CANGA	Q6ft63	candida gla	943	23	82.1	640	2	Q7N531	PHOLL	Q7n531	photorhabdu
871	23	82.1	566	1	AMY_STRGR		P30270	streptomyce	944	23	82.1	641	2	Q5B2V6	EMENI	Q5b2v6	aspergillus
872	23	82.1	566	1	AMY_STRLM		P09794	streptomyce	945	23	82.1	643	2	Q5N031	SYNP6	Q5n031	synecococ
873	23	82.1	566	1	UVRC_AQUAE		O67887	aquifex aeo	946	23	82.1	645	2	Q4LI1X2	9BURK	Q4lix2	burkholderi
874	23	82.1	568	2	Q5HPR6	STAEP	Q5hpr6	staphylococ	947	23	82.1	645	2	Q5FNM9	GLUOX	Q5fnm9	gluconobact
875	23	82.1	568	2	Q8CST0	STAEP	Q8cst0	staphylococ	948	23	82.1	650	2	Q5LTN7	SILPO	Q5ltn7	silicibacte
876	23	82.1	572	2	Q9USN2	MYZPE	Q9usn2	myzus persi	949	23	82.1	651	2	Q5BBC9	EMENI	Q5bbc9	aspergillus
877	23	82.1	573	2	Q7Q270	ANOGA	Q7q270	anopheles g	950	23	82.1	652	2	Q9VS56	DROME	Q9v556	drosophila
878	23	82.1	574	2	Q6Q7P2	ECOLI	Q6q7p2	escherichia	951	23	82.1	654	2	Q4WJ23	ASPFU	Q4wj23	aspergillus
879	23	82.1	574	2	Q573I8	9BACT	Q573i8	uncultured	952	23	82.1	654	2	Q712J2	RAT	Q712j2	rattus norv
880	23	82.1	575	1	ENV_SMRVH		P21412	squirrel mo	953	23	82.1	656	2	Q7V1I9	PROMP	Q7v1i9	prochloroco
881	23	82.1	575	1	IUCA_ECOLI		Q47316	escherichia	954	23	82.1	661	2	Q7VX63	BORPE	Q7vx63	bordetella
882	23	82.1	579	1	UVRC_METTH		O26541	methanobact	955	23	82.1	661	2	Q7W829	BORPA	Q7w829	bordetella
883	23	82.1	580	2	Q4WBV3	ASPFU	Q4wbv3	aspergillus	956	23	82.1	661	2	Q7WLH6	BORBR	Q7wlh6	bordetella
884	23	82.1	581	2	O46219	9DIPT	O46219	culicoides	957	23	82.1	663	2	Q5AVX6	EMENI	Q5avx6	aspergillus
885	23	82.1	588	2	Q83R46	SHIFL	Q83r46	shigella fl	958	23	82.1	666	2	Q8IQ82	DROME	Q8iq82	drosophila
886	23	82.1	590	2	Q610B6	CAEBR	Q610b6	caenorhabdi	959	23	82.1	666	2	Q5Z230	NOCFA	Q5z230	nocardia fa
887	23	82.1	591	2	Q4WFN9	ASPFU	Q4wfn9	aspergillus	960	23	82.1	667	2	Q7V7L7	PROMM	Q7v7l7	prochloroco
888	23	82.1	592	2	Q92Y36	RHIME	Q92y36	rhizobium m	961	23	82.1	669	1	UVRB_XYLFA		Q9per1	xylella fas
889	23	82.1	593	2	Q6KD48	ECOLI	Q6kd48	escherichia	962	23	82.1	669	1	UVRB_XYLFT		Q87at6	xylella fas
890	23	82.1	593	2	Q93PE6	SHIBO	Q93pe6	shigella bo	963	23	82.1	669	2	Q7N6Q1	PHOLL	Q7n6q1	photorhabdu
891	23	82.1	593	2	Q9XCH4	SHIFL	Q9xch4	shigella fl	964	23	82.1	672	2	Q5WNM3	CAEBR	Q5wnm3	caenorhabdi
892	23	82.1	594	1	UL93_HCMVA		P16799	human cytom	965	23	82.1	673	1	UVRB_XANAC		Q8pjb1	caenothomonas
893	23	82.1	594	2	Q4KAL6	PSEF5	Q4kal6	pseudomonas	966	23	82.1	673	1	UVRB_XANCP		Q8p7x1	xanthomonas
894	23	82.1	594	2	Q6RXE6	HCMV	Q6rx66	human cytom	967	23	82.1	673	2	Q4UW79	XANCP	Q4uw79	xanthomonas
895	23	82.1	594	2	Q80MY0	HCMV	Q80my0	human cytom	968	23	82.1	673	2	Q9RBK2	XANCP	Q9rbk2	xanthomonas
896	23	82.1	594	2	Q6SW50	HCMV	Q6sw50	human cytom	969	23	82.1	674	2	O62140	CAEEL	O62140	caenorhabdi
897	23	82.1	595	2	Q8FDV6	ECOL6	Q8fdv6	escherichia	970	23	82.1	677	2	Q8IQ81	DROME	Q8iq81	drosophila
898	23	82.1	597	2	Q5R6I9	PONPY	Q5r6i9	pongo pygma	971	23	82.1	677	2	Q72BN0	DESVH	Q72bn0	desulfovibr
899	23	82.1	597	2	Q4HIV4	CAMLA	Q4hiv4	campylobact	972	23	82.1	678	2	Q5H218	XANOR	Q5h218	xanthomonas
900	23	82.1	598	2	Q55FN1	DICDI	Q55fn1	dictyosteli	973	23	82.1	678	2	Q4UXG7	XANCP	Q4uxg7	xanthomonas
901	23	82.1	598	2	Q8GHE7	SALTY	Q8ghe7	salmonella	974	23	82.1	678	2	Q8P6P5	XANCP	Q8p6p5	xanthomonas
902	23	82.1	598	2	Q939L2	CITFR	Q939l2	citrobacter	975	23	82.1	678	2	Q8PHZ9	XANAC	Q8phz9	xanthomonas
903	23	82.1	598	2	Q57JS0	SALCH	Q57js0	salmonella	976	23	82.1	679	2	O62137	CAEEL	O62137	caenorhabdi
904	23	82.1	598	2	Q9XDP2	ENTAM	Q9xdp2	enterobacte	977	23	82.1	679	2	Q8K1G9	MOUSE	Q8k1g9	mus musculu
905	23	82.1	598	2	Q5PMU8	SALPA	Q5pmu8	salmonella	978	23	82.1	687	2	Q7VCW1	PROMA	Q7vcw1	prochloroco
906	23	82.1	598	2	Q8Z3N6	SALTI	Q8z3n6	salmonella	979	23	82.1	688	2	Q5GXX4	XANOR	Q5gxx4	xanthomonas
907	23	82.1	600	2	Q51EP9	ENTHI	Q51ep9	entamoeba h	980	23	82.1	689	1	ACOX1	_YARLI	O74934	yarrowia li

981 23 82.1 691 2 Q8IQ80 DROME Q8iq80 drosophila
982 23 82.1 693 2 Q4P2J2_USTMA Q4p2j2 ustilago ma
983 23 82.1 693 2 Q87D39_XYLFT Q87d39 xylella fas
984 23 82.1 693 2 Q9PC32_XYLFA Q9pc32 xylella fas
985 23 82.1 694 2 Q4UDG8_THEAN Q4udg8 theileria a
986 23 82.1 695 1 ELF3 ARATH Q82804 arabidopsis
987 23 82.1 697 2 Q4N630_THEPA Q4n630 theileria p
988 23 82.1 697 2 Q8QVE8_9REOV Q8qve8 eyach virus
989 23 82.1 700 1 ACOX3_YARLI Q74936 yarrowia li
990 23 82.1 702 2 Q4SKH2_TETNG Q4skh2 tetraodon n
991 23 82.1 705 2 Q9Y8D5_9PLEO Q9y8d5 cochliobolu
992 23 82.1 708 2 Q7QDU7_ANOGA Q7qdu7 anopheles g
993 23 82.1 718 2 Q5JVVZ5_HUMAN Q5jvz5 homo sapien
994 23 82.1 718 2 Q5GMG3_MOUSE Q5gmg3 mus musculu
995 23 82.1 719 1 ACOX_PICPA Q9y7b1 pichia past
996 23 82.1 720 1 ELMO2_HUMAN Q96jj3 homo sapien
997 23 82.1 720 2 Q5RCC1_PONPY Q5rcc1 pongo pygma
998 23 82.1 720 2 Q66IJ1_XENTR Q66ij1 xenopus tro
999 23 82.1 723 2 Q9DUC4_9VIRU Q9duc4 torque teno
1000 23 82.1 725 2 Q6G471_BARHE Q6g471 bartonella

ALIGNMENTS

RESULT 1
Q6MGX7_BDEBA
ID Q6MGX7_BDEBA PRELIMINARY; PRT; 232 AA.
AC Q6MGX7;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Bd3792;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842656; CAE81152.1; -; Genomic_DNA.
DR InterPro; IPR003730; DUF152.
DR Pfam; PF02578; DUF152; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 232 AA; 25776 MW; 80A2229E03ED2677 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 232;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 148 VRDQILSSL 156

RESULT 2
Q7QAB5_ANOGA
ID Q7QAB5_ANOGA PRELIMINARY; PRT; 455 AA.
AC Q7QAB5;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ENSANGP0000013373 (Fragment).
GN ORFNames=ENSANGG0000010884;
OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBBJ databases.
CC -i- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008898; EAA09186.2; -; Genomic_DNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0016071; P:mRNA metabolism; IEA.
DR InterPro; IPR001313; Pumilio/Puf.
DR SMART; SM00025; Pumilio; 4.
DR PROSITE; PS50303; PUM_HD; 1.
FT NON_TER 1
FT NON_TER 455 455
SQ SEQUENCE 455 AA; 51607 MW; 2D6DD0915F9AC2D1 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 455;
Best Local Similarity 55.6%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 110 IRDQIATSL 118

RESULT 3
Q9ZSH5_ARATH
ID Q9ZSH5_ARATH PRELIMINARY; PRT; 543 AA.
AC Q9ZSH5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE T9E19.1 protein.
GN Names=T9E19.1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA WashU;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Stromatt C., Johnson D., Le T.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AF104920; AAC72877.1; -; Genomic_DNA.
DR PIR; T02022; T02022.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 543 AA; 60444 MW; 8D2438A5F022DBF1 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 543;

Best Local Similarity 66.7%; Pred. No. 5.2e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 6; Conservative 0;

QY 1 VRDQIXXXL 9
Db 69 VRDQIRAAAL 77

RESULT 4
Q9M180 ARATH
ID Q9M180 ARATH PRELIMINARY; PRT; 1113 AA.
AC Q9M180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T5C2_90.
GN Name=T5C2_90;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL138664; CAB81580.1; -; Genomic_DNA.
DR PIR; T47381; T47381.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 1113 AA; 123579 MW; B826A0588B97DDBE CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1113;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 488 VRDQIRAAAL 496

RESULT 5
Q9LK04 ARATH
ID Q9LK04 ARATH PRELIMINARY; PRT; 1119 AA.
AC Q9LK04;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gb|AAC24188.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000382; BAB02940.1; -; Genomic_DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1119 AA; 125821 MW; CD3483A918B08895 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1119;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 488 VRDQIRAAAL 496

RESULT 6
Q9S9Z6 ARATH
ID Q9S9Z6 ARATH PRELIMINARY; PRT; 1132 AA.
AC Q9S9Z6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE F21H2.5.
GN Name=F21H2.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C.,
RA Rowley D., Shinn P., Walker M., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Theologis;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007894; AAD46015.1; -; Genomic_DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48;
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1132 AA; 126285 MW; 7F04A71FE0BF6A2E CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1132;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 468 VRDQIRAAAL 476

RESULT 7
Q9C6N9 ARATH
ID Q9C6N9 ARATH PRELIMINARY; PRT; 1198 AA.
AC Q9C6N9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F28L5.2.
GN Name=F28L5.2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079280; AAG50582.1; -; Genomic_DNA.
DR PIR; E86402; E86402.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase C48.
DR Pfam; PF02902; Peptidase C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1198 AA; 133579 MW; D608AE577CE21339 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1198;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 486 VRDQIRATL 494

RESULT 8
Q9C607_ARATH PRELIMINARY; PRT; 1201 AA.
AC Q9C607;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F14G11.4.
GN Name=F14G11.4;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084221; AAG50521.1; -; Genomic_DNA.
DR PIR; F86386; F86386.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase C48.
DR Pfam; PF02902; Peptidase C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1201 AA; 133995 MW; D10F98847289CFB7 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1201;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 455 VRDQIRAAAL 463

RESULT 9
Q9ZU21_ARATH PRELIMINARY; PRT; 1305 AA.
AC Q9ZU21;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F5F19.8 protein.
GN Name=F5F19.8;
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Lenz C., Liu S.,
RA Li J., Kremenetskaia I., Luros J., Gonzalez A., Altafi H., Araujo R.,
RA Buehler E., Conway A.B., Conn L., Dunn P., Hansen N., Huizar L.,
RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006216; AAD12676.1; -; Genomic_DNA.
DR PIR; H96559; H96559.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase C48.
DR Pfam; PF02902; Peptidase C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1305 AA; 146079 MW; D3C036ABEE2EFD6C CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1305;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 491 VRDQIRAAAL 499

RESULT 10
Q9LTG8_ARATH PRELIMINARY; PRT; 1444 AA.
AC Q9LTG8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similarity to mutator-like transposase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025605; BAA98062.1; -; Genomic_DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase C48.
DR Pfam; PF02902; Peptidase C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1444 AA; 161024 MW; AF0CB23AAAC4EB8B CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1444;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 488 VRDQIRAAAL 496

RESULT 11
Q9SFY0_ARATH PRELIMINARY; PRT; 1444 AA.
AC Q9SFY0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T22C5.24.

DE F21E10.9 protein.
GN Name=F21E10.9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA WashU;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Davidson S., Rohlfing T., David M., O'Brian D.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF058914; AAC13589.1; -; Genomic_DNA.
DR PIR; T01194; T01194.
DR HSSP; P11746; 1MNM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF00319; SRF-TF; 1.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS0066; MADS_BOX_2; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 211 AA; 24251 MW; FEE93738B384E60D CRC64;

Query Match 92.9%; Score 26; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 187 VRDQIQITL 195

RESULT 15
Q7XJK7 ARATH PRELIMINARY; PRT; 218 AA.
AC Q7XJK7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MADS-box protein AGL35.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Siliques;
RA Parenicova L., de Folter S., Kieffer M., Horner D.S., Favalli C.,
RA Busscher J., Cook H.E., Ingram R.M., Kater M.M., Davies B.,
RA Angenent G.C., Colombo L.;
RT "Molecular and Phylogenetic Analyses of the Complete MADS-Box
Transcription Factor Family in Arabidopsis: New Openings to the MADS
World.";
RL Plant Cell 0:0-0(2003).
DR EMBL; AY141246; AAN52810.1; -; mRNA.
DR HSSP; P11831; 1HBX.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF00319; SRF-TF; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS50066; MADS_BOX_2; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 218 AA; 25123 MW; 3D8DF0B56CDBCDD4 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 218;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 194 VRDQIQITL 202

RESULT 16
Q5WH83 BACSK PRELIMINARY; PRT; 272 AA.
AC Q5WH83;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Phosphomethylpyrimidine kinase (EC 2.7.4.7).
GN OrderedLocusNames=ABC1737;
OS Bacillus clausii (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=66692;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-K16;
RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA Kawai S., Ito S., Horikoshi K.;
RT "The complete genome sequence of the alkaliphilic Bacillus clausii
KSM-K16.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006627; BAD64272.1; -; Genomic DNA.
DR GO; GO:0008972; F:phosphomethylpyrimidine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009228; P:thiamin biosynthesis; IEA.
DR InterPro; IPR004399; HMP-P_kinase.
DR InterPro; IPR011611; PfkB_region.
DR Pfam; PF00294; PfkB; 1.
DR TIGRFAMs; TIGR00097; HMP-P_kinase; 1.
KW Complete proteome; Kinase; Transferase.
SQ SEQUENCE 272 AA; 28197 MW; A81DA7A12A298AF7 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 272;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 60 VRDQIEAVL 68

RESULT 17
Q7NKK0 GLOVI PRELIMINARY; PRT; 334 AA.
AC Q7NKK0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gll1477 protein.
GN OrderedLocusNames=gll1477;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC89418.1; -; Genomic_DNA.
DR InterPro; IPR007296; DUF403_bac.
DR Pfam; PF04168; DUF403; 1.
KW Complete proteome.
SQ SEQUENCE 334 AA; 38221 MW; 44519A598FBE1920 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 334;
Best Local Similarity 55.6%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 100 IRDQISSEL 108

RESULT 18
Q8Z110_SALTI PRELIMINARY; PRT; 357 AA.
AC Q8Z110; Q7C426;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Protein kinase.
GN OrderedLocusNames=STY4823, t4520;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627283; CAD06945.1; -; Genomic DNA.
DR EMBL; AE016849; AA071965.1; -; Genomic_DNA.
DR HSSP; P71584; 1MRU.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Transferase.
SQ SEQUENCE 357 AA; 40433 MW; 5C234A0F141516D6 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 255 VRDQIAKYL 263

RESULT 19
Q8MYV0_DROME PRELIMINARY; PRT; 486 AA.
AC Q8MYV0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RH42429p.
GN ORFNames=CG7433;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
DR EMBL; AY113591; AAM29596.1; -; mRNA.
DR HSSP; P80147; 1OHV.
DR Ensembl; CG7433; Drosophila melanogaster.
DR FlyBase; FBgn0036927; CG7433.
DR GO; GO:0003867; F:4-aminobutyrate transaminase activity; IEA.
DR GO; GO:0030170; F:pyridoxal phosphate binding; IEA.
DR GO; GO:0009448; P:gamma-aminobutyric acid metabolism; IEA.
DR InterPro; IPR005814; Aminotrans_3.
DR InterPro; IPR004631; GABAtrns_euk.
DR PANTHER; PTHR11986; Aminotrans_3; 1.
DR Pfam; PF00202; Aminotran_3; 1.
DR TIGRFAMS; TIGR00699; GABAtrns_euk; 1.
KW Pyridoxal phosphate.
SQ SEQUENCE 486 AA; 54530 MW; 7A1C8F1140388D78 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
:|||||
Db 436 VRDQIIGAL 444

RESULT 20
Q9VW68_DROME PRELIMINARY; PRT; 486 AA.
AC Q9VW68;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG7433-PA, isoform A (CG7433-pb, isoform b).
GN Name=CG7433; ORFNames=CG7433;
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genome perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- INTERACTION:
CC Q9VWR8-1:CG6659; NbExp=1; IntAct=EBI-86087, EBI-281215;
CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
CC EMBL; AE003515; AAF49081.1; -; Genomic_DNA.
CC HSP; P80147; IOHV.
CC IntAct; Q9VW68; -.
CC Ensembl; CG7433; Drosophila melanogaster.
CC FlyBase; FBgn0036927; CG7433.
CC GO; GO:0003867; F:4-aminobutyrate transaminase activity; IEA.
CC GO; GO:0030170; F:pyridoxal phosphate binding; IEA.
CC GO; GO:0009448; P:gamma-aminobutyric acid metabolism; IEA.
CC InterPro; IPR005814; Aminotrans_3.
CC InterPro; IPR004631; GABATrans_euk.
CC PANTHER; PTHR11986; Aminotrans_3; 1.
CC Pfam; PF00202; Aminotran_3; 1.
CC TIGRFAMs; TIGR00699; GABATrans_euk; 1.
KW Pyridoxal phosphate.
SQ SEQUENCE 486 AA; 54584 MW; 405BF910A543A51B CRC64;

Query Match 92.9%; Score 26; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 436 VRDQIIIGAL 444

RESULT 21
Q4PCC1 USTMA PRELIMINARY; PRT; 516 AA.
AC Q4PCC1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM02242.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Biye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,

RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP0100079; EAK83364.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 516 AA; 54638 MW; 2BB0AP2853AE4B65 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 516;
Best Local Similarity 66.7%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 382 VRDQITNHL 390

RESULT 22
Q4Q7Y6 LEIMA
ID Q4Q7Y6_LEIMA PRELIMINARY; PRT; 594 AA.
AC Q4Q7Y6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Gamma-tubulin complex subunit, putative.
GN ORFNames=lmjF28.2760;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005266; CAJ05734.1; -; Genomic_DNA.
SQ SEQUENCE 594 AA; 68118 MW; 566DB8C1434C721B CRC64;

Query Match 92.9%; Score 26; DB 2; Length 594;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

Db 320 VRDQVQSAL 328
RESULT 23
Q89S89 BRAJA
ID Q89S89 BRAJA PRELIMINARY; PRT; 642 AA.
AC Q89S89;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B112516 protein.
GN OrderedLocusNames=b112516;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC47781.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 642 AA; 68447 MW; A3F527A4F4323CFF CRC64;

Query Match 92.9%; Score 26; DB 2; Length 642;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 416 VRDQIEADL 424

RESULT 24
Q5LX00 SILPO
ID Q5LX00_SILPO PRELIMINARY; PRT; 743 AA.
AC Q5LX00;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SPO0204;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment.";
RL Nature 432:910-913(2004).
DR EMBL; CP000031; AAV93530.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 743 AA; 81225 MW; 00B63B9D70C629BE CRC64;

Query Match 92.9%; Score 26; DB 2; Length 743;

Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 731 VRDQIDSL 739

RESULT 25

Q7UYC9_RHOBA
ID Q7UYC9_RHOBA PRELIMINARY; PRT; 896 AA.
AC Q7UYC9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable chromosome segregation protein.
GN Ordered locus names=RB714;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294134; CAD71715.1; -; Genomic_DNA.
DR GO; GO:0020037; F:heme binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR012282; Cytochrome_c_r.
DR InterPro; IPR011444; DUF1549.
DR InterPro; IPR011429; PSC3.
DR Pfam; PF07635; PSCytl; 1.
DR Pfam; PF07583; PSCyt2; 1.
DR Pfam; PF07587; PSD1; 1.
KW Complete proteome.
SQ SEQUENCE 896 AA; 101882 MW; 6C5627C8EEA742CE CRC64;

Query Match 92.9%; Score 26; DB 2; Length 896;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 306 VRDQIAGDL 314

RESULT 26

Q65353_NPVAC
ID Q65353_NPVAC PRELIMINARY; PRT; 1084 AA.
AC Q65353;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF B.
OS Autographa californica nuclear polyhedrosis virus (ACMPNV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=mutant FP-D;
RX MEDLINE=90258898; PubMed=1692964;
RA Friesen P.D., Nissen M.S.;
RT "Gene organization and transcription of TED, a lepidopteran
RT retrotansposon integrated within the baculovirus genome."
RL Mol. Cell. Biol. 10:3067-3077 (1990).
DR EMBL; M32662; AAA92249.1; -; Genomic_DNA.

DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
SQ SEQUENCE 1084 AA; 125586 MW; 77E8D0063F83F296 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 154 VRDQITKML 162

RESULT 27

Q4P3H5_USTMA
ID Q4P3H5_USTMA PRELIMINARY; PRT; 1105 AA.
AC Q4P3H5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05338.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA Mccarthy M., Mcdonough S., MCGhee T., Meldrim J., Meneus L., Moru K.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okaowo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,

RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000192; EAK86587.1; -; Genomic_DNA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR003163; Yeast_DNA_bd.
DR Pfam; PF00023; Ank; 3.
DR Pfam; PF02292; APSES; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR ANK repeat; DNA-binding; Hypothetical protein; Repeat.
KW SEQUENCE 1105 AA; 115924 MW; CDCBA6DAABE398DF CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1105;
Best Local Similarity 55.6%; Pred. No. 2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
DB 814 VRDQLSSTL 822

RESULT 28
Q4I1R0 GIBZE
ID Q4I1R0_GIBZE PRELIMINARY; PRT; 1191 AA.
AC Q4I1R0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG08848.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PH-1;
RC Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

-!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000358; EAA72022.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1191 AA; 135430 MW; A09A78310DC82C38 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1191;
Best Local Similarity 55.6%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
DB 804 IRDQILSAL 812

RESULT 29
Q9M0W7 ARATH
ID Q9M0W7_ARATH PRELIMINARY; PRT; 1312 AA.
AC Q9M0W7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein AT4g05280.
GN Name=AT4g05280;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
[1]
RN NUCLEOTIDE SEQUENCE.
RP Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RESULT 30
Q61Z65 CAEBR
ID Q61Z65_CAEBR PRELIMINARY; PRT; 1407 AA.
AC Q61Z65;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG03251 (Fragment).
GN Name=CBG03251;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
[1]
RN NUCLEOTIDE SEQUENCE.
RP The C.briggsae Sequencing Consortium;
RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; CAAC01000012; CAB59791.1; -; Genomic_DNA.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.

Query Match 92.9%; Score 26; DB 2; Length 1312;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
DB 434 VRDQIRAVL 442

RESULT 30
Q61Z65 CAEBR
ID Q61Z65_CAEBR PRELIMINARY; PRT; 1407 AA.
AC Q61Z65;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG03251 (Fragment).
GN Name=CBG03251;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
[1]
RN NUCLEOTIDE SEQUENCE.
RP The C.briggsae Sequencing Consortium;
RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; CAAC01000012; CAB59791.1; -; Genomic_DNA.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.

DR GO: GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO: GO:0006512; P:ubiquitin cycle; IEA.
DR GO: GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001394; Peptidase_C19.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
KW Hypothetical protein.
FT NON_TER 1 1407
FT NON_TER 1407 1407
SQ SEQUENCE 1407 AA; 164607 MW; 139EC18FC9677E4A CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1407;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 135 VRDQIQSVL 143

RESULT 31
Q51WD5 MAGGR
ID Q51WD5_MAGGR PRELIMINARY; PRT; 2240 AA.
AC Q51WD5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG03850.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook K., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gherre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysasselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,

RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001031; EAA50091.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 2240 AA; 255489 MW; 5E00710960191B59 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 2240;
Best Local Similarity 66.7%; Pred. No. 4.2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
Db 617 VRDQIKAKL 625

RESULT 32
Q76YV2 9CAUD
ID Q76YV2_9CAUD PRELIMINARY; PRT; 89 AA.
AC Q76YV2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AehlORF129c;
OS Bacteriophage Aehl.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=227470;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Petrov V., Nolan J., Bertrand C., Letarov A.V., Kriisch H.M.,
RA Karam J.D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Petrov V., Nolan J., Karam J.D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY266303; AAQ17794.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
KW Hypothetical protein.
SQ SEQUENCE 89 AA; 10144 MW; 572457907A444EA8 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 89;
Best Local Similarity 44.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
: | | | :
Db 58 IRDQVAAKL 66

RESULT 33
Q6RXD6 HCMV
ID Q6RXD6_HCMV PRELIMINARY; PRT; 102 AA.
AC Q6RXD6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UL101.
GN Name=UL101;
OS Human cytomegalovirus (HHV-5) (Human herpesvirus 5).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Toledo;
RA Brondke H., Schmitz B., Shenk T., Doerfler W.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY486477; AAS48987.1; -; Genomic DNA.
SQ SEQUENCE 102 AA; 10795 MW; 0CFFFF4C288E02B57 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 102;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 60 VRDQLSSQL 68

RESULT 34
UL101_HCMVA STANDARD; PRT; 115 AA.
ID UL101_HCMVA
AC P16826;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein UL101.
GN Name=UL101;
OS Human cytomegalovirus (strain AD169) (HHV-5) (Human herpesvirus 5).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X17403; CAA35337.1; -; Genomic_DNA.
DR PIR; S09866; S09866.
KW Hypothetical protein.
SQ SEQUENCE 115 AA; 12184 MW; 5D3BDC55D48D904F CRC64;

Query Match 89.3%; Score 25; DB 1; Length 115;
Best Local Similarity 55.6%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 73 VRDQLSSQL 81

RESULT 35
Q58J60_9ACTO PRELIMINARY; PRT; 130 AA.
ID Q58J60_9ACTO
AC Q58J60;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)

DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Putative regulatory protein.
OS Streptomyces noursei ATCC 11455.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=316284;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 11455;
RX PubMed=15990252; DOI=10.1016/j.femsle.2005.05.052;
RA Volokhan O., Sletta H., Sekurova O.N., Ellingsen T.E., Zotchev S.B.;
RT "An unexpected role for the putative 4'-phosphopantetheinyl
RT transferase-encoding gene nysF in the regulation of nystatin
RT biosynthesis in Streptomyces noursei ATCC 11455."
RL FEMS Microbiol. Lett. 249:57-64(2005).
DR EMBL; AY942707; AAX37275.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS00552; HTH_MERR_1; 1.
DR PROSITE; PS50937; HTH_MERR_2; 1.
KW DNA-binding.
SQ SEQUENCE 130 AA; 14496 MW; AE8EF07611855BDC CRC64;

Query Match 89.3%; Score 25; DB 2; Length 130;
Best Local Similarity 55.6%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | : |
Db 103 VRDQVNTQL 111

RESULT 36
Q4XMZ2_PLACH PRELIMINARY; PRT; 139 AA.
ID Q4XMZ2_PLACH
AC Q4XMZ2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC000779.04.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAJ01004543; CAH81720.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 16249 MW; 5A31ABF4B835B1AE CRC64;

Query Match 89.3%; Score 25; DB 2; Length 139;
Best Local Similarity 55.6%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY      1 VRDQIXXXL 9
Db      96 IRDQILLAL 104

RESULT 37
Q9I6T7_PSEAE PRELIMINARY;      PRT;      193 AA.
AC      Q9I6T7;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hypothetical protein.
GN      OrderedLocusNames=PA0201;
OS      Pseudomonas aeruginosa.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=287;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=ATCC 15692 / PA01; PubMed=10984043; DOI=10.1038/35023079;
RX      MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA      Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA      Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA      Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA      Brody R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA      Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA      Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT      "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT      opportunistic pathogen.";
RL      Nature 406:959-964 (2000).
DR      EMBL; AB004458; AAG03590.1; -; Genomic_DNA.
DR      PIR; D83620; D83620.
DR      InterPro; IPR010662; DUF1234.
DR      Pfam; PF06821; DUF1234; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE      193 AA; 21734 MW; 1FDC933C4DD794BE CRC64;

      Query Match      89.3%; Score 25; DB 2; Length 193;
      Best Local Similarity      55.6%; Pred. No. 5.6e+02;
      Matches      5; Conservative      1; Mismatches      3; Indels      0; Gaps      0;

QY      1 VRDQIXXXL 9
Db      87 VRDQVRGAL 95

RESULT 38
Q6SL06_BACLD
ID      Q6SL06_BACLD PRELIMINARY;      PRT;      225 AA.
AC      Q6SL06; Q62WF3;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Hypothetical DNA-binding protein, putative transcriptional
DE      regulator.
GN      OrderedLocusNames=BL03804, BLi01354;
OS      Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=279010;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=15383718; DOI=10.1159/000079829;
RA      Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA      Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA      Ehrenreich A., Gottschalk G.;
RT      "The complete genome sequence of Bacillus licheniformis DSM13, an
RT      organism with great industrial potential.";
RL      J. Mol. Microbiol. Biotechnol. 7:204-211 (2004).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
```

```
RA      Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA      Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA      Olesen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA      Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA      Ehrlich S.D., Berka R.M.;
RT      "Complete genome sequence of the industrial bacterium Bacillus
RT      licheniformis and comparisons with closely related Bacillus species.";
RL      Genome Biol. 5:RESEARCH077.1-RESEARCH077.12 (2004).
DR      EMBL; AE017333; AAU40258.1; -; Genomic_DNA.
DR      EMBL; CP000002; AAU22905.1; -; Genomic_DNA.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      GO; GO:0006350; P:transcription; IEA.
DR      InterPro; IPR011711; GntR_C.
DR      InterPro; IPR000524; HTH_GntR.
DR      Pfam; PF07729; FCD; 1.
DR      Pfam; PF00392; GntR; 1.
DR      PRINTS; PR00035; HTHGNTR.
DR      SMART; SM00345; HTH_GNTR; 1.
DR      PROSITE; PS50949; HTH_GNTR; 1.
KW      Complete proteome; DNA-binding; Hypothetical protein; Transcription;
KW      Transcription regulation.
SQ      SEQUENCE      225 AA; 26425 MW; 7F60B5D7770477EB CRC64;

      Query Match      89.3%; Score 25; DB 2; Length 225;
      Best Local Similarity      66.7%; Pred. No. 6.6e+02;
      Matches      6; Conservative      0; Mismatches      3; Indels      0; Gaps      0;

QY      1 VRDQIXXXL 9
Db      11 VRDQIYQVL 19

RESULT 39
Q6VT06_9VIRU
ID      Q6VT06_9VIRU PRELIMINARY;      PRT;      229 AA.
AC      Q6VT06;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Hypothetical protein orf11C.
GN      Name=orf11C;
OS      Vibrio parahaemolyticus phage VP16C.
OC      Viruses.
OX      NCBI_TaxID=238893;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=14563879; DOI=10.1128/JB.185.21.6434-6447.2003;
RA      Seguritan V., Feng I.W., Rohwer F., Swift M., Segall A.M.;
RT      "Genome sequences of two closely related Vibrio parahaemolyticus
RT      phages, VP16T and VP16C.";
RL      J. Bacteriol. 185:6434-6447 (2003).
DR      EMBL; AY328853; AAQ96542.1; -; Genomic_DNA.
KW      Hypothetical protein.
SQ      SEQUENCE      229 AA; 25407 MW; 38CFF17DC93CD5DA CRC64;

      Query Match      89.3%; Score 25; DB 2; Length 229;
      Best Local Similarity      55.6%; Pred. No. 6.8e+02;
      Matches      5; Conservative      1; Mismatches      3; Indels      0; Gaps      0;

QY      1 VRDQIXXXL 9
Db      18 VRDQVAQIL 26

RESULT 40
P90912_CAEEL
ID      P90912_CAEEL PRELIMINARY;      PRT;      243 AA.
AC      P90912;
DT      01-MAY-1997 (TrEMBLrel. 03, Created)
DT      01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```

DE Hypothetical protein K07A1.7.
GN ORFNames=K07A1.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2; PubMed=9851916;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81097; CAB03168.1; -; Genomic_DNA.
DR PIR; T23381; T23381.
DR Ensembl; K07A1.7; Caenorhabditis elegans.
DR WormBase; WBGene00010614; K07A1.7.
DR WormPep; K07A1.7; CE11852.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 243 AA; 27965 MW; 38042BB7692A2988 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 243;
Best Local Similarity 55.6%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 206 VRDQLATQL 214
||||: |

RESULT 41
EVG1_DROME
ID EVG1_DROME STANDARD; PRT; 250 AA.
AC Q9VSS7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0193 protein EVG1 homolog.
GN ORFNames=CG5280;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., P.A.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- INTERACTION:
CC Q9VIS0:CG10466; NbExp=1; IntAct=EBI-95286, EBI-171642;
CC Q8IR89:CG11727; NbExp=1; IntAct=EBI-95286, EBI-131736;
CC Q9W4B4:CG5062; NbExp=1; IntAct=EBI-95286, EBI-121808;
CC Q9VHR7:CG7352; NbExp=1; IntAct=EBI-95286, EBI-184070;
CC Q9W4Y8:EG:BACH7M4.4; NbExp=1; IntAct=EBI-95286, EBI-153177;
CC -!- SIMILARITY: Belongs to the UPF0193 (EVG1) family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE003553; AAF50335.1; -; Genomic_DNA.
DR IntAct; Q9VSS7; -.
DR Ensembl; CG5280; Drosophila melanogaster.
DR FlyBase; FBgn0035952; CG5280.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR007914; UPF0193.
DR Pfam; PF05250; UPF0193; 1.
KW Hypothetical protein.
SQ SEQUENCE 250 AA; 28564 MW; 67A8D1FAF2E72252 CRC64;

Query Match 89.3%; Score 25; DB 1; Length 250;
Best Local Similarity 55.6%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 220 IRDQIAERL 228
:|||||
|

RESULT 42
Q8MZ83_DROME
ID Q8MZ83_DROME PRELIMINARY; PRT; 250 AA.
AC Q8MZ83;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE AT27353p.
GN ORFNames=CG5280;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.


```
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY113308; AAM29313.1; -; mRNA.
DR Ensembl; CG5280; Drosophila melanogaster.
DR FlyBase; FBgn0035952; CG5280.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR007914; UPF0193.
DR Pfam; PF05250; UPF0193; 1.
SQ SEQUENCE 250 AA; 28564 MW; E5A8D1FAF2E7225C CRC64;

Query Match 89.3%; Score 25; DB 2; Length 250;
Best Local Similarity 55.6%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 220 IRDQIAERL 228
:|||||
|

RESULT 43
Q72MR9 LEPI
ID Q72MR9_LEPIC PRELIMINARY; PRT; 291 AA.
AC Q72MR9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LiC13123;
OS Leptospiira interrogans (serogroup Icterohaemorrhagiae / serovar Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospiira.
OC NCBI_TaxID=44275;
[1]
RP NUCLEOTIDE SEQUENCE.
STRAIN=Flocruz L1-130;
PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorry H.,
Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Giglioti E.A.,
Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,
Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,
Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
"Comparative genomics of two Leptospiira interrogans serovars reveals
novel insights into physiology and pathogenesis.";
J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AE017299; AAS71669.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR PRINTS; PR01590; HTHFIS.
KW Complete proteome.
SQ SEQUENCE 291 AA; 33139 MW; 28CA78664F554E1D CRC64;

Query Match 89.3%; Score 25; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 266 VRDQIRKDL 274
:|||||
|

RESULT 44
Q8EZE2 LEPIN
ID Q8EZE2_LEPIN PRELIMINARY; PRT; 291 AA.
AC Q8EZE2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA3912;
OS Leptospiira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospiira.
OC NCBI_TaxID=173;
[1]
RP NUCLEOTIDE SEQUENCE.
STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
Xu J.-G., Zhao G.-P.;
"Unique physiological and pathogenic features of Leptospiira
interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011546; AAN51110.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011989; ARM-like.
DR InterPro; IPR002197; HTH_Fis.
DR PRINTS; PR01590; HTHFIS.
KW Complete proteome.
SQ SEQUENCE 291 AA; 33125 MW; E11EE8639BBEDE15 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 266 VRDQIRKDL 274
:|||||
|

RESULT 45
Q8G822 BIFLO
ID Q8G822_BIFLO PRELIMINARY; PRT; 296 AA.
AC Q8G822;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Possible TetR-type transcriptional regulator.
GN OrderedLocusNames=BL0061;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OC NCBI_TaxID=216816;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
STRAIN=NCC 2705;
MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
Pridmore R.D., Arigoni F.;
"The genome sequence of Bifidobacterium longum reflects its adaptation
to the human gastrointestinal tract.";
Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014295; AAN23927.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PROSITE; PS50977; HTH_TETR_2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33740 MW; 4522D8B92DE048D6 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 296;
Best Local Similarity 66.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 199 VRDQIEDDL 207
|||||

RESULT 46
PERN_ARMRU STANDARD; PRT; 327 AA.
ID PERN_ARMRU
AC Q42517;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Peroxidase N precursor (EC 1.11.1.7) (Neutral peroxidase).
GN Name=HRPN;
OS Armoracia rusticana (Horseradish) (Armoracia laphatifolia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Armoracia.
OX NCBI_TaxID=3704;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Root;
RX MEDLINE=91159476; PubMed=2001399;
RA Bartonek-Roxa E., Eriksson H., Mattiasson B.;
RT "The cDNA sequence of a neutral horseradish peroxidase.";
RL Biochim. Biophys. Acta 1088:245-250(1991).
CC -!- FUNCTION: Removal of H(2)O(2), oxidation of toxic reductants,
CC biosynthesis and degradation of lignin, suberization, auxin
CC catabolism, response to environmental stresses such as wounding,
CC pathogen attack and oxidative stress. These functions might be
CC dependent on each isozyme/isoform in each plant tissue.
CC -!- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.
CC -!- COFACTOR: Binds 2 calcium ions per subunit.
CC -!- COFACTOR: Binds 1 heme B (iron-protoporphyrin IX) group per
CC subunit.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the peroxidase family. Classical plant
CC (class III) peroxidase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X57564; CAA40796.1; -; mRNA.
CC PIR; S14268; S14268.
CC HSSP; Q39034; 1QGJ.
CC SMR; Q42517; 29-327.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR000823; Plant_peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PRINTS; PR00461; PLPEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
DR PROSITE; PS50873; PEROXIDASE_4; 1.
KW Calcium; Glycoprotein; Heme; Hydrogen peroxide; Iron; Metal-binding;
KW Multigene family; Oxidoreductase; Peroxidase;
KW Pyrrolidone carboxylic acid; Signal.
FT SIGNAL 1 28 Potential.
FT CHAIN 29 327 Peroxidase N.

FT ACT_SITE 70 70 Proton acceptor (By similarity).
FT METAL 71 71 Calcium 1 (By similarity).
FT METAL 74 74 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 76 76 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 78 78 Calcium 1 (By similarity).
FT METAL 80 80 Calcium 1 (By similarity).
FT METAL 193 193 Iron (heme axial ligand) (By similarity).
FT METAL 194 194 Calcium 2 (By similarity).
FT METAL 245 245 Calcium 2 (By similarity).
FT METAL 248 248 Calcium 2 (By similarity).
FT METAL 253 253 Calcium 2 (By similarity).
FT BINDING 163 163 Substrate (By similarity).
FT SITE 66 66 Transition state stabilizer (By
FT similarity).
FT MOD_RES 29 29 Pyrrolidone carboxylic acid (By
FT similarity).
FT CARBOHYD 155 155 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 182 182 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 209 209 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 239 239 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 247 247 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 281 281 N-linked (GlcNAc . .) (Potential).
FT DISULFID 39 116 By similarity.
FT DISULFID 72 77 By similarity.
FT DISULFID 122 323 By similarity.
FT DISULFID 200 232 By similarity.
SQ SEQUENCE 327 AA; 35126 MW; 5C427EBDD0A2CCDF CRC64;

Query Match 89.3%; Score 25; DB 1; Length 327;
Best Local Similarity 55.6%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 46 VRDQVKIAL 54
|||||

RESULT 47

Q61Y40 CAEBR
ID Q61Y40_CAEBR PRELIMINARY; PRT; 358 AA.
AC Q61Y40;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG03720.
GN Name=CBG03720;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAC01000016; CAE60166.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 358 AA; 40922 MW; 07DFE0D15316381C CRC64;

Query Match 89.3%; Score 25; DB 2; Length 358;
Best Local Similarity 55.6%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 309 VRDQLATQL 317
|||||

RESULT 48

Q4WFK8 ASPFU
ID Q4WFK8 ASPFU PRELIMINARY; PRT; 422 AA.
AC Q4WFK8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Afu3G01870;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penalba M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF01000010; EAL86469.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 422 AA; 48117 MW; 2149AA2F3C7EBBDA CRC64;

Query Match 89.3%; Score 25; DB 2; Length 422;
Best Local Similarity 55.6%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 241 VRDEISTTL 249

RESULT 49
Q8D8W3 VIBVU
ID Q8D8W3 VIBVU PRELIMINARY; PRT; 437 AA.
AC Q8D8W3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=VV12853;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE016806; AAO1189.1; -; Genomic DNA.
DR GO; GO:0042597; C:periplasmic space; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR010869; DUF1501.
DR InterPro; IPR006311; Tat.
DR Pfam; PF07394; DUF1501; 1.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW Complete proteome.
SQ SEQUENCE 437 AA; 47709 MW; 6250B6240B304753 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 437;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 243 IRDQLSAAL 251

RESULT 50
Q7MLK9 VIBVY
ID Q7MLK9 VIBVY PRELIMINARY; PRT; 437 AA.
AC Q7MLK9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VV1418.
GN OrderedLocusNames=VV1418;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC94182.1; -; Genomic_DNA.
DR InterPro; IPR010869; DUF1501.
DR InterPro; IPR006311; Tat.
DR Pfam; PF07394; DUF1501; 1.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 437 AA; 47755 MW; 117B4F65E97D43E8 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 437;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
Db 243 IRDQLSAAL 251

Search completed: May 12, 2006, 10:52:58
Job time : 90.6923 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 43.5897 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-34
Perfect score: 34
Sequence: 1 TLREWL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query Match			Length	DB ID	Description
	Score	Match	Length			
1	34	100.0	6	3	AAB16978	Aab16978 TPO-mimet
2	34	100.0	6	5	ABB72864	Abb72864 TPO mimet
3	34	100.0	6	6	ABG71746	Abg71746 TPO recep
4	34	100.0	6	7	ADJ73015	Adj73015 TPO mimet
5	34	100.0	6	8	ADJ52650	Adj52650 CH1 delet
6	34	100.0	6	8	ADJ51611	Adj51611 CH1 delet
7	34	100.0	9	7	ADJ73040	Adj73040 TPO mimet
8	34	100.0	9	7	ADJ73037	Adj73037 TPO mimet
9	34	100.0	9	7	ADJ73041	Adj73041 TPO mimet
10	34	100.0	9	7	ADJ73039	Adj73039 TPO mimet
11	34	100.0	9	7	ADJ73038	Adj73038 TPO mimet
12	34	100.0	10	2	AAW09471	Aaw09471 Thrombopo
13	34	100.0	10	2	AAW09566	Aaw09566 Thrombopo
14	34	100.0	10	2	AAW36622	Aaw36622 Thrombopo
15	34	100.0	10	2	AAW36717	Aaw36717 Thrombopo
16	34	100.0	10	3	AAB17000	Aab17000 TPO-mimet
17	34	100.0	10	4	AAU25841	Aau25841 Human thr
18	34	100.0	10	4	AAU25936	Aau25936 Human thr
19	34	100.0	10	5	ABB72886	Abb72886 TPO mimet
20	34	100.0	10	8	ADJ52672	Adj52672 CH1 delet
21	34	100.0	10	8	ADJ51633	Adj51633 CH1 delet
22	34	100.0	11	3	AAB17001	Aab17001 TPO-mimet
23	34	100.0	11	5	ABB72887	Abb72887 TPO mimet
24	34	100.0	11	8	ADJ52673	Adj52673 CH1 delet

25	34	100.0	11	8	ADJ51634	Adj51634 CH1 delet
26	34	100.0	12	2	AAW09476	Aaw09476 Thrombopo
27	34	100.0	12	2	AAW09581	Aaw09581 Thrombopo
28	34	100.0	12	2	AAW09577	Aaw09577 Thrombopo
29	34	100.0	12	2	AAW09524	Aaw09524 Thrombopo
30	34	100.0	12	2	AAW36728	Aaw36728 Thrombopo
31	34	100.0	12	2	AAW36627	Aaw36627 Thrombopo
32	34	100.0	12	2	AAW36675	Aaw36675 Thrombopo
33	34	100.0	12	3	AAB17002	Aab17002 TPO-mimet
34	34	100.0	12	3	AAB16993	Aab16993 TPO-mimet
35	34	100.0	12	4	AAU25846	Aau25846 Human thr
36	34	100.0	12	4	AAU25951	Aau25951 Human thr
37	34	100.0	12	4	AAU25894	Aau25894 Human thr
38	34	100.0	12	4	AAU25947	Aau25947 Human thr
39	34	100.0	12	5	ABB72888	Abb72888 TPO mimet
40	34	100.0	12	5	ABB72879	Abb72879 TPO mimet
41	34	100.0	12	7	ADJ73030	Adj73030 TPO mimet
42	34	100.0	12	8	ADJ52665	Adj52665 CH1 delet
43	34	100.0	12	8	ADJ52674	Adj52674 CH1 delet
44	34	100.0	12	8	ADJ51626	Adj51626 CH1 delet
45	34	100.0	12	8	ADJ51635	Adj51635 CH1 delet
46	34	100.0	13	2	AAW36732	Aaw36732 Thrombopo
47	34	100.0	13	3	AAB17015	Aab17015 TPO-mimet
48	34	100.0	13	3	AAB17003	Aab17003 TPO-mimet
49	34	100.0	13	5	ABB72901	Abb72901 TPO mimet
50	34	100.0	13	5	ABB72889	Abb72889 TPO mimet
51	34	100.0	13	7	ADJ73054	Adj73054 TPO mimet
52	34	100.0	13	7	ADJ73052	Adj73052 TPO mimet
53	34	100.0	13	7	ADJ73056	Adj73056 TPO mimet
54	34	100.0	13	7	ADJ73053	Adj73053 TPO mimet
55	34	100.0	13	7	ADJ73055	Adj73055 TPO mimet
56	34	100.0	13	8	ADJ52687	Adj52687 CH1 delet
57	34	100.0	13	8	ADJ52675	Adj52675 CH1 delet
58	34	100.0	13	8	ADJ51636	Adj51636 CH1 delet
59	34	100.0	13	8	ADJ51648	Adj51648 CH1 delet
60	34	100.0	14	3	AAB17017	Aab17017 TPO-mimet
61	34	100.0	14	3	AAB17016	Aab17016 TPO-mimet
62	34	100.0	14	3	AAB16969	Aab16969 TPO-mimet
63	34	100.0	14	3	AAB17004	Aab17004 TPO-mimet
64	34	100.0	14	5	ABB72903	Abb72903 TPO mimet
65	34	100.0	14	5	ABB72855	Abb72855 TPO mimet
66	34	100.0	14	5	ABB72890	Abb72890 TPO mimet
67	34	100.0	14	5	ABB72902	Abb72902 TPO mimet
68	34	100.0	14	7	ADJ73005	Adj73005 TPO mimet
69	34	100.0	14	8	ADJ52676	Adj52676 CH1 delet
70	34	100.0	14	8	ADJ52689	Adj52689 CH1 delet
71	34	100.0	14	8	ADJ52640	Adj52640 CH1 delet
72	34	100.0	14	8	ADJ52688	Adj52688 CH1 delet
73	34	100.0	14	8	ADJ51601	Adj51601 CH1 delet
74	34	100.0	14	8	ADJ51650	Adj51650 CH1 delet
75	34	100.0	14	8	ADJ51649	Adj51649 CH1 delet
76	34	100.0	14	8	ADJ51637	Adj51637 CH1 delet
77	34	100.0	15	3	AAB17018	Aab17018 TPO-mimet
78	34	100.0	15	5	ABB72904	Abb72904 TPO mimet
79	34	100.0	15	8	ADJ52691	Adj52691 CH1 delet
80	34	100.0	15	8	ADJ52690	Adj52690 CH1 delet
81	34	100.0	15	8	ADJ51652	Adj51652 CH1 delet
82	34	100.0	15	8	ADJ51651	Adj51651 CH1 delet
83	34	100.0	16	2	AAW09464	Aaw09464 Thrombopo
84	34	100.0	16	2	AAW33329	Aaw33329 Thrombopo
85	34	100.0	16	3	AAB17019	Aab17019 TPO-mimet
86	34	100.0	16	4	AAU25829	Aau25829 Human thr
87	34	100.0	16	5	ABB72905	Abb72905 TPO mimet
88	34	100.0	16	7	ADJ73057	Adj73057 TPO mimet
89	34	100.0	16	8	ADJ52692	Adj52692 CH1 delet
90	34	100.0	16	8	ADJ51653	Adj51653 CH1 delet
91	34	100.0	18	2	AAW09460	Aaw09460 Thrombopo
92	34	100.0	18	2	AAW09583	Aaw09583 Thrombopo
93	34	100.0	18	2	AAW09498	Aaw09498 Thrombopo
94	34	100.0	18	2	AAW09486	Aaw09486 Thrombopo
95	34	100.0	18	2	AAW36734	Aaw36734 Thrombopo
96	34	100.0	18	2	AAW36637	Aaw36637 Thrombopo
97	34	100.0	18	2	AAW36649	Aaw36649 Thrombopo

98	34	100.0	18	2	AAW33027	Aaw33027 Thrombopo	171	34	100.0	422	3	AAy96815	Aay96815 A. thalia
99	34	100.0	18	2	AAW36652	Aaw36652 Thrombopo	172	34	100.0	455	3	AGS9392	Aag59392 Arabidops
100	34	100.0	18	3	AAB17026	Aab17026 TPO-mimet	173	34	100.0	486	3	AGS9391	Aag59391 Arabidops
101	34	100.0	18	4	AAU25856	Aau25856 Human thr	174	34	100.0	496	9	ABM91663	Abm91663 M. xanthu
102	34	100.0	18	4	AAU25835	Aau25835 Human thr	175	34	100.0	530	3	AGS9390	Aag59390 Arabidops
103	34	100.0	18	4	AAU25868	Aau25868 Human thr	176	32	94.1	9	2	AAW35429	Aaw35429 Thrombopo
104	34	100.0	18	4	AAU25953	Aau25953 Human thr	177	32	94.1	9	4	AAU26003	Aau26003 Human thr
105	34	100.0	18	4	AAU25837	Aau25837 Human thr	178	32	94.1	10	2	AAW35427	Aaw35427 Thrombopo
106	34	100.0	18	4	AAU25824	Aau25824 Human thr	179	32	94.1	10	4	AAU26002	Aau26002 Human thr
107	34	100.0	18	4	AAU25871	Aau25871 Human thr	180	32	94.1	11	2	AAW35425	Aaw35425 Thrombopo
108	34	100.0	18	5	ABB72912	Abb72912 TPO mimet	181	32	94.1	11	4	AAU26001	Aau26001 Human thr
109	34	100.0	18	7	ADJ73064	Adj73064 TPO mimet	182	32	94.1	12	2	AAW35423	Aaw35423 Thrombopo
110	34	100.0	18	7	ADN59664	Adn59664 Thrombopo	183	32	94.1	12	4	AAU26000	Aau26000 Human thr
111	34	100.0	18	7	ADN59672	Adn59672 Thrombopo	184	32	94.1	12	8	ADM72530	Adm72530 TPO mimet
112	34	100.0	18	7	ADN59658	Adn59658 Thrombopo	185	32	94.1	13	2	AAW09467	Aaw09467 Thrombopo
113	34	100.0	18	7	ADN59675	Adn59675 Thrombopo	186	32	94.1	13	2	AAW35399	Aaw35399 Thrombopo
114	34	100.0	18	7	ADN59680	Adn59680 Thrombopo	187	32	94.1	13	2	AAW35404	Aaw35404 Thrombopo
115	34	100.0	18	7	ADN59668	Adn59668 Thrombopo	188	32	94.1	13	2	AAW35417	Aaw35417 Thrombopo
116	34	100.0	18	7	ADN59669	Adn59669 Thrombopo	189	32	94.1	13	2	AAW33033	Aaw33033 Thrombopo
117	34	100.0	18	7	ADN59657	Adn59657 Thrombopo	190	32	94.1	13	2	AAW35413	Aaw35413 Thrombopo
118	34	100.0	18	7	ADN59673	Adn59673 Thrombopo	191	32	94.1	13	2	AAW35406	Aaw35406 Thrombopo
119	34	100.0	18	7	ADN59667	Adn59667 Thrombopo	192	32	94.1	13	2	AAW35405	Aaw35405 Thrombopo
120	34	100.0	18	8	ADJ52699	Adj52699 CH1 delet	193	32	94.1	13	2	AAW35422	Aaw35422 Thrombopo
121	34	100.0	18	8	ADJ51660	Adj51660 CH1 delet	194	32	94.1	13	2	AAW35397	Aaw35397 Thrombopo
122	34	100.0	19	4	AAU25836	Aau25836 Human thr	195	32	94.1	13	4	AAU26039	Aau26039 Human thr
123	34	100.0	22	7	ADN59840	Adn59840 TMP pepti	196	32	94.1	13	4	AAU26041	Aau26041 Human thr
124	34	100.0	22	7	ADN59824	Adn59824 TMP pepti	197	32	94.1	13	4	AAU25994	Aau25994 Human thr
125	34	100.0	22	7	ADN59836	Adn59836 TMP pepti	198	32	94.1	13	4	AAU25997	Aau25997 Human thr
126	34	100.0	22	7	ADN59835	Adn59835 TMP pepti	199	32	94.1	13	4	AAU25991	Aau25991 Human thr
127	34	100.0	22	7	ADN59831	Adn59831 TMP pepti	200	32	94.1	13	4	AAU25990	Aau25990 Human thr
128	34	100.0	22	7	ADN59825	Adn59825 TMP pepti	201	32	94.1	13	4	AAU25984	Aau25984 Human thr
129	34	100.0	22	7	ADN59839	Adn59839 TMP pepti	202	32	94.1	13	8	ADM72529	Adm72529 TPO mimet
130	34	100.0	22	7	ADN59834	Adn59834 TMP pepti	203	32	94.1	13	8	ADM72528	Adm72528 TPO mimet
131	34	100.0	23	7	ADN59775	Adn59775 Peptide-v	204	32	94.1	14	2	AAW09466	Aaw09466 Thrombopo
132	34	100.0	23	7	ADN59778	Adn59778 Peptide-v	205	32	94.1	14	2	AAW09462	Aaw09462 Thrombopo
133	34	100.0	23	7	ADN59793	Adn59793 Peptide-v	206	32	94.1	14	2	AAW09465	Aaw09465 Thrombopo
134	34	100.0	23	7	ADN59796	Adn59796 Peptide-v	207	32	94.1	14	2	AAW09482	Aaw09482 Thrombopo
135	34	100.0	25	7	ADN59726	Adn59726 Thrombopo	208	32	94.1	14	2	AAW33031	Aaw33031 Thrombopo
136	34	100.0	25	7	ADN59696	Adn59696 Thrombopo	209	32	94.1	14	2	AAW35412	Aaw35412 Thrombopo
137	34	100.0	25	7	ADN59730	Adn59730 Thrombopo	210	32	94.1	14	2	AAW35407	Aaw35407 Thrombopo
138	34	100.0	25	7	ADN59710	Adn59710 Thrombopo	211	32	94.1	14	2	AAW35408	Aaw35408 Thrombopo
139	34	100.0	25	7	ADN59698	Adn59698 Thrombopo	212	32	94.1	14	2	AAW35398	Aaw35398 Thrombopo
140	34	100.0	25	7	ADN59716	Adn59716 Thrombopo	213	32	94.1	14	2	AAW36633	Aaw36633 Thrombopo
141	34	100.0	25	7	ADN59718	Adn59718 Thrombopo	214	32	94.1	14	2	AAW33029	Aaw33029 Thrombopo
142	34	100.0	25	7	ADN59740	Adn59740 Thrombopo	215	32	94.1	14	2	AAW35396	Aaw35396 Thrombopo
143	34	100.0	25	7	ADN59720	Adn59720 Thrombopo	216	32	94.1	14	2	AAW35401	Aaw35401 Thrombopo
144	34	100.0	25	7	ADN59744	Adn59744 Thrombopo	217	32	94.1	14	2	AAW35403	Aaw35403 Thrombopo
145	34	100.0	36	7	ADN59763	Adn59763 Peptide-v	218	32	94.1	14	2	AAW36647	Aaw36647 Thrombopo
146	34	100.0	36	7	ADN59766	Adn59766 Peptide-v	219	32	94.1	14	2	AAW35400	Aaw35400 Thrombopo
147	34	100.0	36	7	ADN59756	Adn59756 Peptide-v	220	32	94.1	14	2	AAW35402	Aaw35402 Thrombopo
148	34	100.0	41	7	ADN59769	Adn59769 Peptide-v	221	32	94.1	14	2	AAW33032	Aaw33032 Thrombopo
149	34	100.0	41	7	ADN59816	Adn59816 Peptide-	222	32	94.1	14	2	AAW66732	Aaw66732 Peptide c
150	34	100.0	41	7	ADN59772	Adn59772 Peptide-v	223	32	94.1	14	3	AAB17014	Aab17014 TPO-mimet
151	34	100.0	42	7	ADN59818	Adn59818 Peptide-	224	32	94.1	14	4	AAU25987	Aau25987 Human thr
152	34	100.0	46	7	ADN59781	Adn59781 Peptide-v	225	32	94.1	14	4	AAU25826	Aau25826 Human thr
153	34	100.0	46	7	ADN59787	Adn59787 Peptide-v	226	32	94.1	14	4	AAU25993	Aau25993 Human thr
154	34	100.0	46	7	ADN59790	Adn59790 Peptide-v	227	32	94.1	14	4	AAU25852	Aau25852 Human thr
155	34	100.0	46	7	ADN59784	Adn59784 Peptide-v	228	32	94.1	14	4	AAU25989	Aau25989 Human thr
156	34	100.0	160	4	ABB58297	Abb58297 Drosophil	229	32	94.1	14	4	AAU25983	Aau25983 Human thr
157	34	100.0	226	3	AAy58613	Aay58613 Protein r	230	32	94.1	14	4	AAU25985	Aau25985 Human thr
158	34	100.0	235	3	AAG57785	Aag57785 Arabidops	231	32	94.1	14	4	AAU25995	Aau25995 Human thr
159	34	100.0	235	3	AAG57786	Aag57786 Arabidops	232	32	94.1	14	4	AAU25992	Aau25992 Human thr
160	34	100.0	235	3	AAG55516	Aag55516 Arabidops	233	32	94.1	14	4	AAU25986	Aau25986 Human thr
161	34	100.0	235	3	AAG57791	Aag57791 Arabidops	234	32	94.1	14	4	AAU26040	Aau26040 Human thr
162	34	100.0	243	7	ADJ69556	Adj69556 Human hea	235	32	94.1	14	4	AAU25866	Aau25866 Human thr
163	34	100.0	302	3	AAG55515	Aag55515 Arabidops	236	32	94.1	14	4	AAU25988	Aau25988 Human thr
164	34	100.0	304	7	ADF07784	Adf07784 Bacterial	237	32	94.1	14	5	ABB72900	Abb72900 TPO mimet
165	34	100.0	313	6	ABU89792	Abu89792 Novel hum	238	32	94.1	14	7	ADJ73051	Adj73051 TPO mimet
166	34	100.0	313	8	ADO09332	Ado09332 Novel hum	239	32	94.1	14	8	ADJ52686	Adj52686 CH1 delet
167	34	100.0	325	7	ADB75420	Adb75420 Prostate	240	32	94.1	14	8	ADJ51647	Adj51647 CH1 delet
168	34	100.0	326	5	ABP43878	Abp43878 LRP16 pro	241	32	94.1	18	2	AAW09596	Aaw09596 Thrombopo
169	34	100.0	408	3	AAG57790	Aag57790 Arabidops	242	32	94.1	18	2	AAW09456	Aaw09456 Thrombopo
170	34	100.0	413	3	AAG57789	Aag57789 Arabidops	243	32	94.1	18	2	AAW33023	Aaw33023 Thrombopo

244	32	94.1	18	2	AAW36747	Aaw36747	Thrombopo	317	31	91.2	13	2	AAW36779	Aaw36779	Thrombopo
245	32	94.1	18	2	AAW36748	Aaw36748	Thrombopo	318	31	91.2	13	2	AAW36783	Aaw36783	Thrombopo
246	32	94.1	18	3	AAB17020	Aab17020	TPO-mimet	319	31	91.2	13	3	AAB16995	Aab16995	TPO-mimet
247	32	94.1	18	4	AAU25820	Aau25820	Human thr	320	31	91.2	13	4	AAU26018	Aau26018	Human thr
248	32	94.1	18	4	AAU25966	Aau25966	Human thr	321	31	91.2	13	4	AAU26008	Aau26008	Human thr
249	32	94.1	18	4	AAU25967	Aau25967	Human thr	322	31	91.2	13	4	AAU26012	Aau26012	Human thr
250	32	94.1	18	5	ABB72906	Abb72906	TPO mimet	323	31	91.2	13	4	AAU26015	Aau26015	Human thr
251	32	94.1	18	7	ADJ73058	Adj73058	TPO mimet	324	31	91.2	13	4	AAU26035	Aau26035	Human thr
252	32	94.1	18	8	ADJ52693	Adj52693	CH1 delet	325	31	91.2	13	5	ABB72881	Abb72881	TPO mimet
253	32	94.1	18	8	ADJ51654	Adj51654	CH1 delet	326	31	91.2	13	7	ADJ73003	Adj73003	TPO mimet
254	32	94.1	19	2	AAW09458	Aaw09458	Thrombopo	327	31	91.2	13	7	ADJ73032	Adj73032	TPO mimet
255	32	94.1	19	2	AAW33025	Aaw33025	Thrombopo	328	31	91.2	13	8	ADJ52667	Adj52667	CH1 delet
256	32	94.1	19	2	AAU25822	Aau25822	Human thr	329	31	91.2	13	8	ADJ51599	Adj51599	CH1 delet
257	32	94.1	19	4	AAU26042	Aau26042	Human thr	330	31	91.2	13	8	ADJ51628	Adj51628	CH1 delet
258	32	94.1	25	4	AAU26042	Aau26042	Human thr	331	31	91.2	13	8	ADJ51628	Adj51628	CH1 delet
259	32	94.1	25	8	ADM72531	Adm72531	TPO mimet	332	31	91.2	13	8	ADM72525	Adm72525	TPO mimet
260	32	94.1	109	3	AAB58250	Aab58250	Lung canc	333	31	91.2	13	8	ADM72488	Adm72488	TPO mimet
261	32	94.1	169	7	ABM86829	Abm86829	Rice abio	334	31	91.2	13	8	ADM72489	Adm72489	TPO mimet
262	32	94.1	176	5	AAE25747	Aae25747	Rice FT h	335	31	91.2	13	8	ADM72524	Adm72524	TPO mimet
263	32	94.1	176	8	ADR04263	Adr04263	Rice FT h	336	31	91.2	14	2	AAW09463	Aaw09463	Thrombopo
264	32	94.1	313	3	AAG45687	Aag45687	Arabidops	337	31	91.2	14	2	AAW09468	Aaw09468	Thrombopo
265	32	94.1	313	3	AAG17697	Aag17697	Arabidops	338	31	91.2	14	2	AAW09541	Aaw09541	Thrombopo
266	32	94.1	318	8	ADY22882	Ady22882	Plant ful	339	31	91.2	14	2	AAW33030	Aaw33030	Thrombopo
267	32	94.1	413	8	ADX89692	Adx89692	Plant ful	340	31	91.2	14	2	AAW33034	Aaw33034	Thrombopo
268	32	94.1	450	3	AAG50584	Aag50584	Arabidops	341	31	91.2	14	2	AAW36773	Aaw36773	Thrombopo
269	32	94.1	475	3	AAG50583	Aag50583	Arabidops	342	31	91.2	14	2	AAW36770	Aaw36770	Thrombopo
270	32	94.1	499	3	AAG50582	Aag50582	Arabidops	343	31	91.2	14	2	AAW36782	Aaw36782	Thrombopo
271	32	94.1	569	3	AAG41863	Aag41863	Arabidops	344	31	91.2	14	2	AAW36692	Aaw36692	Thrombopo
272	32	94.1	570	8	ADX90628	Adx90628	Plant ful	345	31	91.2	14	2	AAW36774	Aaw36774	Thrombopo
273	32	94.1	577	3	AAG41861	Aag41861	Arabidops	346	31	91.2	14	2	AAW36788	Aaw36788	Thrombopo
274	32	94.1	834	6	ABU44553	Abu44553	Protein e	347	31	91.2	14	2	AAW66715	Aaw66715	Peptide c
275	32	94.1	834	9	AEB91502	Aeb91502	Microbial	348	31	91.2	14	2	AAW66730	Aaw66730	Peptide c
276	31	91.2	8	2	AAW33037	Aaw33037	Thrombopo	349	31	91.2	14	2	ADI24843	Adi24843	AF 12505
277	31	91.2	8	4	AAU25982	Aau25982	Human thr	350	31	91.2	14	2	AAY96515	Aay96515	Thrombopo
278	31	91.2	8	5	ABP51678	Abp51678	TPO mimet	351	31	91.2	14	3	AAB16962	Aab16962	TPO-mimet
279	31	91.2	8	8	ADQ16692	Adq16692	TPO mimet	352	31	91.2	14	3	AAB16968	Aab16968	TPO-mimet
280	31	91.2	8	9	ADV44430	Adv44430	Agonist T	353	31	91.2	14	3	AAB16999	Aab16999	TPO-mimet
281	31	91.2	8	9	AEB12902	Aeb12902	TPO mimet	354	31	91.2	14	4	AAU26009	Aau26009	Human thr
282	31	91.2	10	2	AAW09473	Aaw09473	Thrombopo	355	31	91.2	14	4	AAU26013	Aau26013	Human thr
283	31	91.2	10	2	AAW09507	Aaw09507	Thrombopo	356	31	91.2	14	4	AAU25911	Aau25911	Human thr
284	31	91.2	10	2	AAW36624	Aaw36624	Thrombopo	357	31	91.2	14	4	AAU26006	Aau26006	Human thr
285	31	91.2	10	2	AAW36658	Aaw36658	Thrombopo	358	31	91.2	14	4	AAU25827	Aau25827	Human thr
286	31	91.2	10	3	AAB16990	Aab16990	TPO-mimet	359	31	91.2	14	4	AAU26010	Aau26010	Human thr
287	31	91.2	10	3	AAB17006	Aab17006	TPO-mimet	360	31	91.2	14	4	AAU26019	Aau26019	Human thr
288	31	91.2	10	4	AAU25843	Aau25843	Human thr	361	31	91.2	14	4	AAU26036	Aau26036	Human thr
289	31	91.2	10	4	AAU25877	Aau25877	Human thr	362	31	91.2	14	4	AAU26037	Aau26037	Human thr
290	31	91.2	10	5	ABB72876	Abb72876	TPO mimet	363	31	91.2	14	4	AAU26004	Aau26004	Human thr
291	31	91.2	10	5	ABB72892	Abb72892	TPO mimet	364	31	91.2	14	5	ABB72854	Abb72854	TPO mimet
292	31	91.2	10	7	ADJ73027	Adj73027	TPO mimet	365	31	91.2	14	5	ABB72885	Abb72885	TPO mimet
293	31	91.2	10	7	ADJ73043	Adj73043	TPO mimet	366	31	91.2	14	5	ABB72853	Abb72853	TPO mimet
294	31	91.2	10	8	ADJ52678	Adj52678	CH1 delet	367	31	91.2	14	5	ABP51669	Abp51669	Thrombopo
295	31	91.2	10	8	ADJ52662	Adj52662	CH1 delet	368	31	91.2	14	5	AAE18011	Aae18011	Human lig
296	31	91.2	10	8	ADJ51623	Adj51623	CH1 delet	369	31	91.2	14	6	ABG71747	Abg71747	TPO recep
297	31	91.2	10	8	ADJ51639	Adj51639	CH1 delet	370	31	91.2	14	7	ABR62907	Abbr62907	Thrombopo
298	31	91.2	11	2	AAW36791	Aaw36791	Thrombopo	371	31	91.2	14	7	ADC33697	Adc33697	Erythropro
299	31	91.2	12	2	AAW09570	Aaw09570	Thrombopo	372	31	91.2	14	7	ADJ73036	Adj73036	TPO mimet
300	31	91.2	12	2	AAW09568	Aaw09568	Thrombopo	373	31	91.2	14	7	ADJ73004	Adj73004	TPO mimet
301	31	91.2	12	2	AAW09572	Aaw09572	Thrombopo	374	31	91.2	14	7	ADN59652	Adn59652	Thrombopo
302	31	91.2	12	2	AAW36723	Aaw36723	Thrombopo	375	31	91.2	14	8	ADJ52671	Adj52671	CH1 delet
303	31	91.2	12	2	AAW36787	Aaw36787	Thrombopo	376	31	91.2	14	8	ADJ52639	Adj52639	CH1 delet
304	31	91.2	12	2	AAW36719	Aaw36719	Thrombopo	377	31	91.2	14	8	ADJ51632	Adj51632	CH1 delet
305	31	91.2	12	2	AAW36721	Aaw36721	Thrombopo	378	31	91.2	14	8	ADJ51600	Adj51600	CH1 delet
306	31	91.2	12	2	AAW36781	Aaw36781	Thrombopo	379	31	91.2	14	8	ADL27293	Adl27293	Amino aci
307	31	91.2	12	3	AAB16994	Aab16994	TPO-mimet	380	31	91.2	14	8	ADM72503	Adm72503	TPO mimet
308	31	91.2	12	4	AAU25938	Aau25938	Human thr	381	31	91.2	14	8	ADM72483	Adm72483	TPO mimet
309	31	91.2	12	4	AAU25942	Aau25942	Human thr	382	31	91.2	14	8	ADM72526	Adm72526	TPO mimet
310	31	91.2	12	4	AAU25940	Aau25940	Human thr	383	31	91.2	14	8	ADM72495	Adm72495	TPO mimet
311	31	91.2	12	4	AAU26014	Aau26014	Human thr	384	31	91.2	14	8	ADM72487	Adm72487	TPO mimet
312	31	91.2	12	5	ABB72880	Abb72880	TPO mimet	385	31	91.2	14	8	ADM72497	Adm72497	TPO mimet
313	31	91.2	12	7	ADJ73031	Adj73031	TPO mimet	386	31	91.2	14	8	ADM72527	Adm72527	TPO mimet
314	31	91.2	12	8	ADJ52666	Adj52666	CH1 delet	387	31	91.2	14	8	ADQ16584	Adq16584	Agonist T
315	31	91.2	12	8	ADJ51627	Adj51627	CH1 delet	388	31	91.2	14	8	ADT92482	Adt92482	Thrombopo
316	31	91.2	13	2	AAW36792	Aaw36792	Thrombopo	389	31	91.2	14	9	ADU70210	Adu70210	Thrombopo

390	31	91.2	14	9	ADU70206	Adu70206 Thrombopo	463	31	91.2	18	3	AAB16957	Aab16957 PEGylated
391	31	91.2	14	9	ADU75982	Adu75982 Peptide-b	464	31	91.2	18	3	AAB16956	Aab16956 PEGylated
392	31	91.2	14	9	ADU75979	Adu75979 Thrombopo	465	31	91.2	18	3	AAB17024	Aab17024 TPO-mimet
393	31	91.2	14	9	ADV44319	Adv44319 Agonist T	466	31	91.2	18	3	AAB17025	Aab17025 TPO-mimet
394	31	91.2	14	9	ABB12792	Aeb12792 TPO mimet	467	31	91.2	18	4	AAU25872	Aau25872 Human thr
395	31	91.2	15	2	AAW35416	Aaw35416 Thrombopo	468	31	91.2	18	4	AAU25867	Aau25867 Human thr
396	31	91.2	15	2	AAW36784	Aaw36784 Thrombopo	469	31	91.2	18	4	AAU25869	Aau25869 Human thr
397	31	91.2	15	2	AAW36780	Aaw36780 Thrombopo	470	31	91.2	18	4	AAU25823	Aau25823 Human thr
398	31	91.2	15	2	AAW36776	Aaw36776 Thrombopo	471	31	91.2	18	4	AAU25965	Aau25965 Human thr
399	31	91.2	15	2	AAW66731	Aaw66731 Peptide c	472	31	91.2	18	4	AAU25962	Aau25962 Human thr
400	31	91.2	15	2	AAW66717	Aaw66717 Peptide c	473	31	91.2	18	5	ABB72910	Abb72910 TPO mimet
401	31	91.2	15	2	AAW66718	Aaw66718 Peptide c	474	31	91.2	18	5	ABB72911	Abb72911 TPO mimet
402	31	91.2	15	2	AAW66714	Aaw66714 Peptide c	475	31	91.2	18	5	ABP51687	Abp51687 TPO mimet
403	31	91.2	15	2	AAW66721	Aaw66721 Peptide c	476	31	91.2	18	5	ABP51689	Abp51689 TPO mimet
404	31	91.2	15	2	AAW66712	Aaw66712 Peptide c	477	31	91.2	18	5	ABP51688	Abp51688 TPO mimet
405	31	91.2	15	3	AAB20684	Aab20684 Thrombocy	478	31	91.2	18	5	ABP51677	Abp51677 TPO mimet
406	31	91.2	15	4	AAU26022	Aau26022 Human thr	479	31	91.2	18	5	ABP51686	Abp51686 TPO mimet
407	31	91.2	15	4	AAU25996	Aau25996 Human thr	480	31	91.2	18	5	ABP51674	Abp51674 TPO mimet
408	31	91.2	15	4	AAU26026	Aau26026 Human thr	481	31	91.2	18	5	ABP51693	Abp51693 TPO mimet
409	31	91.2	15	4	AAU26011	Aau26011 Human thr	482	31	91.2	18	5	ABP51684	Abp51684 TPO mimet
410	31	91.2	15	4	AAU26020	Aau26020 Human thr	483	31	91.2	18	5	ABP51683	Abp51683 TPO mimet
411	31	91.2	15	4	AAU25831	Aau25831 Human thr	484	31	91.2	18	5	ABP51685	Abp51685 TPO mimet
412	31	91.2	15	4	AAU26023	Aau26023 Human thr	485	31	91.2	18	5	ABP51691	Abp51691 TPO mimet
413	31	91.2	15	4	AAU26007	Aau26007 Human thr	486	31	91.2	18	5	ABP51673	Abp51673 TPO mimet
414	31	91.2	15	4	AAU26038	Aau26038 Human thr	487	31	91.2	18	5	ABP51690	Abp51690 TPO mimet
415	31	91.2	15	5	ABP51670	Abp51670 Thrombopo	488	31	91.2	18	5	ABP51675	Abp51675 TPO mimet
416	31	91.2	15	7	ABR62908	Abr62908 Thrombopo	489	31	91.2	18	5	ABP51692	Abp51692 TPO mimet
417	31	91.2	15	8	ADM72485	Adm72485 TPO mimet	490	31	91.2	18	5	ABP51679	Abp51679 TPO mimet
418	31	91.2	15	8	ADM72479	Adm72479 TPO mimet	491	31	91.2	18	7	ADJ73062	Adj73062 TPO mimet
419	31	91.2	15	8	ADM72502	Adm72502 TPO mimet	492	31	91.2	18	7	ADJ73063	Adj73063 TPO mimet
420	31	91.2	15	8	ADM72492	Adm72492 TPO mimet	493	31	91.2	18	7	ADN59681	Adn59681 Thrombopo
421	31	91.2	15	8	ADM72478	Adm72478 TPO mimet	494	31	91.2	18	7	ADN59660	Adn59660 Thrombopo
422	31	91.2	15	8	ADM72533	Adm72533 TPO mimet	495	31	91.2	18	7	ADN59663	Adn59663 Thrombopo
423	31	91.2	15	8	ADM72496	Adm72496 TPO mimet	496	31	91.2	18	7	ADN59679	Adn59679 Thrombopo
424	31	91.2	15	8	ADM72490	Adm72490 TPO mimet	497	31	91.2	18	7	ADN59654	Adn59654 Thrombopo
425	31	91.2	15	8	ADM72486	Adm72486 TPO mimet	498	31	91.2	18	7	ADN59678	Adn59678 Thrombopo
426	31	91.2	15	8	ADM72491	Adm72491 TPO mimet	499	31	91.2	18	7	ADN59815	Adn59815 Thrombopo
427	31	91.2	15	8	ADM72522	Adm72522 TPO mimet	500	31	91.2	18	7	ADN59661	Adn59661 Thrombopo
428	31	91.2	15	8	ADM72523	Adm72523 TPO mimet	501	31	91.2	18	7	ADN59653	Adn59653 Thrombopo
429	31	91.2	15	8	ADM72493	Adm72493 TPO mimet	502	31	91.2	18	7	ADN59676	Adn59676 Thrombopo
430	31	91.2	15	8	ADM72482	Adm72482 TPO mimet	503	31	91.2	18	7	ADN59812	Adn59812 Thrombopo
431	31	91.2	15	8	ADM72494	Adm72494 TPO mimet	504	31	91.2	18	7	ADN59666	Adn59666 Thrombopo
432	31	91.2	15	8	ADQ16585	Adql6585 TPO mimet	505	31	91.2	18	7	ADN59670	Adn59670 Thrombopo
433	31	91.2	15	8	ADT92483	Adt92483 Modified	506	31	91.2	18	7	ADN59656	Adn59656 Thrombopo
434	31	91.2	15	9	ADU70209	Adu70209 Thrombopo	507	31	91.2	18	8	ADJ52698	Adj52698 CH1 delet
435	31	91.2	15	9	ADU75981	Adu75981 Peptide-b	508	31	91.2	18	8	ADJ52697	Adj52697 CH1 delet
436	31	91.2	15	9	ADV44320	Adv44320 Agonist T	509	31	91.2	18	8	ADJ51658	Adj51658 CH1 delet
437	31	91.2	15	9	ABB12793	Aeb12793 TPO mimet	510	31	91.2	18	8	ADJ51659	Adj51659 CH1 delet
438	31	91.2	16	2	AAW19534	Aaw19534 Thrombopo	511	31	91.2	18	8	ADQ16611	Adql6611 TPO mimet
439	31	91.2	16	2	AAW33035	Aaw33035 Thrombopo	512	31	91.2	18	8	ADQ16619	Adql6619 TPO mimet
440	31	91.2	16	2	AAW36775	Aaw36775 Thrombopo	513	31	91.2	18	8	ADQ16621	Adql6621 TPO mimet
441	31	91.2	16	2	AAW36771	Aaw36771 Thrombopo	514	31	91.2	18	8	ADQ16641	Adql6641 TPO mimet
442	31	91.2	16	2	AAW66709	Aaw66709 Peptide c	515	31	91.2	18	8	ADQ16646	Adql6646 TPO mimet
443	31	91.2	16	2	AAW66713	Aaw66713 Peptide c	516	31	91.2	18	8	ADQ16607	Adql6607 TPO mimet
444	31	91.2	16	2	AAW66733	Aaw66733 Peptide c	517	31	91.2	18	8	ADQ16615	Adql6615 TPO mimet
445	31	91.2	16	2	AAW66716	Aaw66716 Peptide c	518	31	91.2	18	8	ADQ16693	Adql6693 TPO mimet
446	31	91.2	16	4	AAU26021	Aau26021 Human thr	519	31	91.2	18	8	ADQ16627	Adql6627 TPO mimet
447	31	91.2	16	4	AAU26005	Aau26005 Human thr	520	31	91.2	18	8	ADQ16625	Adql6625 TPO mimet
448	31	91.2	16	4	AAU26043	Aau26043 Human thr	521	31	91.2	18	8	ADQ16617	Adql6617 TPO mimet
449	31	91.2	16	4	AAU25832	Aau25832 Human thr	522	31	91.2	18	8	ADQ16629	Adql6629 TPO mimet
450	31	91.2	16	8	ADM72532	Adm72532 TPO mimet	523	31	91.2	18	8	ADQ16613	Adql6613 TPO mimet
451	31	91.2	16	8	ADM72484	Adm72484 TPO mimet	524	31	91.2	18	8	ADQ16623	Adql6623 TPO mimet
452	31	91.2	18	2	AAW09592	Aaw09592 Thrombopo	525	31	91.2	18	8	ADQ16605	Adql6605 TPO mimet
453	31	91.2	18	2	AAW09499	Aaw09499 Thrombopo	526	31	91.2	18	8	ADQ16609	Adql6609 TPO mimet
454	31	91.2	18	2	AAW09595	Aaw09595 Thrombopo	527	31	91.2	18	9	ADV44355	Adv44355 Agonist T
455	31	91.2	18	2	AAW09497	Aaw09497 Thrombopo	528	31	91.2	18	9	ADV44431	Adv44431 Anti-teta
456	31	91.2	18	2	AAW09459	Aaw09459 Thrombopo	529	31	91.2	18	9	ADV44345	Adv44345 Agonist T
457	31	91.2	18	2	AAW36650	Aaw36650 Thrombopo	530	31	91.2	18	9	ADV44351	Adv44351 Agonist T
458	31	91.2	18	2	AAW36746	Aaw36746 Thrombopo	531	31	91.2	18	9	ADV44357	Adv44357 Agonist T
459	31	91.2	18	2	AAW36653	Aaw36653 Thrombopo	532	31	91.2	18	9	ADV44349	Adv44349 Agonist T
460	31	91.2	18	2	AAW33026	Aaw33026 Thrombopo	533	31	91.2	18	9	ADV44343	Adv44343 Agonist T
461	31	91.2	18	2	AAW36743	Aaw36743 Thrombopo	534	31	91.2	18	9	ADV44347	Adv44347 Agonist T
462	31	91.2	18	2	AAW36648	Aaw36648 Thrombopo	535	31	91.2	18	9	ADV44384	Adv44384 Agonist T

536	31	91.2	18	9	ADV44353	Adv44353 Agonist T	609	31	91.2	22	8	ADQ16714	Adq16714 Immunoglo
537	31	91.2	18	9	ADV44363	Adv44363 Agonist T	610	31	91.2	22	8	ADQ16713	Adq16713 Immunoglo
538	31	91.2	18	9	ADV44361	Adv44361 Agonist T	611	31	91.2	22	8	ADQ16709	Adq16709 Immunoglo
539	31	91.2	18	9	ADV44379	Adv44379 Agonist T	612	31	91.2	22	8	ADQ16706	Adq16706 Immunoglo
540	31	91.2	18	9	ADV44344	Adv44344 Agonist T	613	31	91.2	22	8	ADQ16699	Adq16699 TPO mimet
541	31	91.2	18	9	ADV44359	Adv44359 Agonist T	614	31	91.2	22	8	ADQ16712	Adq16712 Immunoglo
542	31	91.2	18	9	ADV44365	Adv44365 Agonist T	615	31	91.2	22	8	ADQ16707	Adq16707 Immunoglo
543	31	91.2	18	9	ADV44367	Adv44367 Agonist T	616	31	91.2	22	8	ADQ16711	Adq16711 Immunoglo
544	31	91.2	18	9	ADV44468	Adv44468 Agonist T	617	31	91.2	22	8	ADQ16708	Adq16708 Immunoglo
545	31	91.2	18	9	AEb12826	Aeb12826 TPO mimet	618	31	91.2	22	8	ADQ16710	Adq16710 Immunoglo
546	31	91.2	18	9	AEb12820	Aeb12820 TPO mimet	619	31	91.2	22	9	ADV44435	Adv44435 Modified
547	31	91.2	18	9	AEb12840	Aeb12840 TPO mimet	620	31	91.2	22	9	ADV44449	Adv44449 Anti-teta
548	31	91.2	18	9	AEb12857	Aeb12857 Antibody	621	31	91.2	22	9	ADV44443	Adv44443 Anti-teta
549	31	91.2	18	9	AEb12834	Aeb12834 TPO mimet	622	31	91.2	22	9	ADV44444	Adv44444 Anti-teta
550	31	91.2	18	9	AEb12830	Aeb12830 TPO mimet	623	31	91.2	22	9	ADV44448	Adv44448 Anti-teta
551	31	91.2	18	9	AEb12939	Aeb12939 TPO mimet	624	31	91.2	22	9	ADV44442	Adv44442 Anti-teta
552	31	91.2	18	9	AEb12832	Aeb12832 TPO mimet	625	31	91.2	22	9	ADV44445	Adv44445 Anti-teta
553	31	91.2	18	9	AEb12852	Aeb12852 TPO mimet	626	31	91.2	22	9	ADV44450	Adv44450 Anti-teta
554	31	91.2	18	9	AEb12816	Aeb12816 TPO mimet	627	31	91.2	22	9	ADV44446	Adv44446 Anti-teta
555	31	91.2	18	9	AEb12818	Aeb12818 TPO mimet	628	31	91.2	22	9	AEb12918	Aeb12918 TPO mimet
556	31	91.2	18	9	AEb12828	Aeb12828 TPO mimet	629	31	91.2	22	9	AEb12917	Aeb12917 TPO mimet
557	31	91.2	18	9	AEb12836	Aeb12836 TPO mimet	630	31	91.2	22	9	AEb12919	Aeb12919 TPO mimet
558	31	91.2	18	9	AEb12903	Aeb12903 Heavy cha	631	31	91.2	22	9	AEb12986	AEb12986 TT antibo
559	31	91.2	18	9	AEb12822	Aeb12822 TPO mimet	632	31	91.2	22	9	AEb12986	Aeb12920 TPO mimet
560	31	91.2	18	9	AEb12824	Aeb12824 TPO mimet	633	31	91.2	22	9	AEb12916	Aeb12916 TPO mimet
561	31	91.2	18	9	AEb12838	Aeb12838 TPO mimet	634	31	91.2	22	9	AEb12914	Aeb12914 TPO mimet
562	31	91.2	19	2	AAw09494	Aaw09494 Thrombopo	635	31	91.2	22	9	AEb12921	Aeb12921 TPO mimet
563	31	91.2	19	2	AAw09461	Aaw09461 Thrombopo	636	31	91.2	22	9	AEb12921	Aeb12921 TPO mimet
564	31	91.2	19	2	AAw09491	Aaw09491 Thrombopo	637	31	91.2	22	9	AEb12907	Aeb12907 TT antibo
565	31	91.2	19	2	AAw09493	Aaw09493 Thrombopo	638	31	91.2	22	9	AEb12915	Aeb12915 TPO mimet
566	31	91.2	19	2	AAw09457	Aaw09457 Thrombopo	639	31	91.2	25	7	ADN59694	Adn59694 Thrombopo
567	31	91.2	19	2	AAw09492	Aaw09492 Thrombopo	640	31	91.2	25	7	ADN59736	Adn59736 Thrombopo
568	31	91.2	19	2	AAw36651	Aaw36651 Thrombopo	641	31	91.2	25	7	ADN59742	Adn59742 Thrombopo
569	31	91.2	19	2	AAw33024	Aaw33024 Thrombopo	642	31	91.2	25	7	ADN59689	Adn59689 Thrombopo
570	31	91.2	19	2	AAw36643	Aaw36643 Thrombopo	643	31	91.2	25	7	ADN59704	Adn59704 Thrombopo
571	31	91.2	19	2	AAw36644	Aaw36644 Thrombopo	644	31	91.2	25	7	ADN59722	Adn59722 Thrombopo
572	31	91.2	19	2	AAw36645	Aaw36645 Thrombopo	645	31	91.2	25	7	ADN59738	Adn59738 Thrombopo
573	31	91.2	19	2	AAw36645	Aaw36645 Thrombopo	646	31	91.2	25	7	ADN59702	Adn59702 Thrombopo
574	31	91.2	19	2	AAw35418	Aaw35418 Thrombopo	647	31	91.2	25	7	ADN59691	Adn59691 Thrombopo
575	31	91.2	19	2	AAw36642	Aaw36642 Thrombopo	648	31	91.2	25	7	ADN59708	Adn59708 Thrombopo
576	31	91.2	19	3	AAb17022	Aab17022 TPO-mimet	649	31	91.2	25	7	ADN59732	Adn59732 Thrombopo
577	31	91.2	19	3	AAb17021	Aab17021 TPO-mimet	650	31	91.2	25	7	ADN59714	Adn59714 Thrombopo
578	31	91.2	19	4	AAu25861	Aau25861 Human thr	651	31	91.2	28	3	AAb17285	Aab17285 TPO-mimet
579	31	91.2	19	4	AAu25864	Aau25864 Human thr	652	31	91.2	28	5	ABp51682	Abp51682 TPO mimet
580	31	91.2	19	4	AAu25863	Aau25863 Human thr	653	31	91.2	28	7	ADJ73013	Adj73013 TPO mimet
581	31	91.2	19	4	AAu25862	Aau25862 Human thr	654	31	91.2	28	8	ADJ52648	Adj52648 CH1 delet
582	31	91.2	19	4	AAu25870	Aau25870 Human thr	655	31	91.2	28	8	ADJ51609	Adj51609 CH1 delet
583	31	91.2	19	4	AAu25825	Aau25825 Human thr	656	31	91.2	28	8	ADQ16636	Adq16636 Tetanus t
584	31	91.2	19	4	AAu25998	Aau25998 Human thr	657	31	91.2	28	9	ADV44374	Adv44374 Modified
585	31	91.2	19	4	AAu25821	Aau25821 Human thr	658	31	91.2	28	9	AEb12847	Aeb12847 Antibody
586	31	91.2	19	5	ABb73391	Abb73391 TPO-mimet	659	31	91.2	29	3	AAb16971	Aab16971 TPO-mimet
587	31	91.2	19	5	ABb72908	Abb72908 TPO mimet	660	31	91.2	29	3	AAb16975	Aab16975 TPO-mimet
588	31	91.2	19	5	ABb72907	Abb72907 TPO mimet	661	31	91.2	29	3	AAb16976	Aab16976 TPO-mimet
589	31	91.2	19	5	ABb73390	Abb73390 TPO-mimet	662	31	91.2	29	3	AAb17286	Aab17286 TPO-mimet
590	31	91.2	19	7	ADJ73060	Adj73060 TPO mimet	663	31	91.2	29	3	AAb16970	Aab16970 TPO-mimet
591	31	91.2	19	7	ADJ73059	Adj73059 TPO mimet	664	31	91.2	29	5	ABb72862	Abb72862 TPO mimet
592	31	91.2	19	8	ADJ52695	Adj52695 CH1 delet	665	31	91.2	29	5	ABb72857	Abb72857 TPO mimet
593	31	91.2	19	8	ADJ52694	Adj52694 CH1 delet	666	31	91.2	29	5	ABb72861	Abb72861 TPO mimet
594	31	91.2	19	8	ADJ51655	Adj51655 CH1 delet	667	31	91.2	29	5	ABb72856	Abb72856 TPO mimet
595	31	91.2	19	8	ADJ51656	Adj51656 CH1 delet	668	31	91.2	29	7	ADJ73011	Adj73011 TPO mimet
596	31	91.2	20	3	AAb18003	Aab18003 Fc-TMP pe	669	31	91.2	29	7	ADJ73007	Adj73007 TPO mimet
597	31	91.2	20	3	AAb17929	Aab17929 TPO-mimet	670	31	91.2	29	7	ADJ73006	Adj73006 TPO mimet
598	31	91.2	20	5	ABb73403	Abb73403 TPO mimet	671	31	91.2	29	8	ADJ52642	Adj52642 CH1 delet
599	31	91.2	21	7	ADN59687	Adn59687 Thrombopo	672	31	91.2	29	8	ADJ52646	Adj52646 CH1 delet
600	31	91.2	22	7	ADN59820	Adn59820 TMP pepti	673	31	91.2	29	8	ADJ51603	Adj51603 CH1 delet
601	31	91.2	22	7	ADN59821	Adn59821 TMP pepti	674	31	91.2	29	8	ADJ51602	Adj51602 CH1 delet
602	31	91.2	22	7	ADN59828	Adn59828 TMP pepti	675	31	91.2	29	8	ADJ51607	Adj51607 CH1 delet
603	31	91.2	22	7	ADN59830	Adn59830 TMP pepti	676	31	91.2	29	8	ADJ51607	Adj51607 CH1 delet
604	31	91.2	22	7	ADN59823	Adn59823 TMP pepti	677	31	91.2	30	3	AAb17287	Aab17287 TPO-mimet
605	31	91.2	22	7	ADN59827	Adn59827 TMP pepti	678	31	91.2	31	3	AAb17288	Aab17288 TPO-mimet
606	31	91.2	22	7	ADN59833	Adn59833 TMP pepti	679	31	91.2	31	3	AAb16974	Aab16974 TPO-mimet
607	31	91.2	22	7	ADN59819	Adn59819 TMP pepti	680	31	91.2	31	3	AAb16973	Aab16973 TPO-mimet
608	31	91.2	22	7	ADN59837	Adn59837 TMP pepti	681	31	91.2	31	5	ABb72860	Abb72860 TPO mimet

682	31	91.2	31	Abb72859	TPO mimet	755	31	91.2	143	6	ABG71750	Abg71750	Antibody
683	31	91.2	31	Adj73009	TPO mimet	756	31	91.2	144	6	ABG71748	Abg71748	Antibody
684	31	91.2	31	Adj73010	TPO mimet	757	31	91.2	150	9	AEB12945	Aeb12945	Antibody
685	31	91.2	31	Adj52644	CH1 delet	758	31	91.2	158	4	AAU29357	Aau29357	Novel mar
686	31	91.2	31	Adj52645	CH1 delet	759	31	91.2	165	9	ABM92612	Abm92612	M. xanthu
687	31	91.2	31	Adj51606	CH1 delet	760	31	91.2	225	8	ADQ16704	Adq16704	Modified
688	31	91.2	31	Adj51605	CH1 delet	761	31	91.2	234	9	AEB12912	Aeb12912	Antibody
689	31	91.2	32	AAy96520	Thrombopo	762	31	91.2	247	3	AAB16958	Aabi6958	Fc-TMP pr
690	31	91.2	32	AAB17289	TPO-mimet	763	31	91.2	247	3	AAB16961	Aabi6961	TMP-Fc pr
691	31	91.2	32	AAB17297	TPO-mimet	764	31	91.2	247	5	ABB73411	Abb73411	Fc-TPO mi
692	31	91.2	33	AAB17290	TPO-mimet	765	31	91.2	247	5	ABB73414	Abb73414	TMP-Fc am
693	31	91.2	34	AAy96527	Thrombopo	766	31	91.2	249	9	ADV44440	Adv44440	pAX116 va
694	31	91.2	34	AAB17291	TPO-mimet	767	31	91.2	268	3	AAB16959	Aabi6959	Fc-TMP-TM
695	31	91.2	35	AAB17292	TPO-mimet	768	31	91.2	268	5	ABB73412	Abb73412	Fc-TMP-TM
696	31	91.2	36	AAy96525	Thrombopo	769	31	91.2	269	3	AAy96531	Aay96531	Human IGG
697	31	91.2	36	AAy96523	Thrombopo	770	31	91.2	269	3	AAB16960	Aabi6960	TMP-TMP-F
698	31	91.2	36	AAy96524	Thrombopo	771	31	91.2	269	5	ABB73413	Abb73413	TMP-TMP-F
699	31	91.2	36	AAy96526	Thrombopo	772	31	91.2	282	8	ADS24597	Ads24597	Bacterial
700	31	91.2	36	AAB17307	TPO-mimet	773	31	91.2	282	9	AEB12930	Aebi2930	Antibody
701	31	91.2	36	AAB17293	TPO-mimet	774	31	91.2	303	4	ABUS3220	Abus3220	Human met
702	31	91.2	36	AAB17303	TPO-mimet	775	31	91.2	308	4	AAG91969	Aag91969	C glutami
703	31	91.2	36	AAB16963	TPO-mimet	776	31	91.2	376	8	ABM84031	Abm84031	Human dia
704	31	91.2	36	AAB17301	TPO-mimet	777	31	91.2	377	8	ABM84030	Abm84030	Human dia
705	31	91.2	36	AAB17306	TPO-mimet	778	31	91.2	404	6	ABP56521	Abp56521	Pseudomon
706	31	91.2	36	ABB72403	TPO-mimet	779	31	91.2	405	2	AAy04128	Aay04128	Pseudomon
707	31	91.2	37	AAB17294	TPO-mimet	780	31	91.2	410	6	ADA66611	Ada66611	Mouse Tra
708	31	91.2	38	AAB17295	TPO-mimet	781	31	91.2	412	2	AAR08264	Aar08264	Tumour gr
709	31	91.2	39	AAB17304	TPO-mimet	782	31	91.2	412	2	AAR20621	Aar20621	Transform
710	31	91.2	39	AAB17305	TPO-mimet	783	31	91.2	412	2	AAR73598	Aar73598	Human TGF
711	31	91.2	40	AAB17302	TPO-mimet	784	31	91.2	412	2	AAW80417	Aaw80417	Amino aci
712	31	91.2	40	ADN59753	Peptide-v	785	31	91.2	412	5	ABB90765	Abb90765	Human Tum
713	31	91.2	40	ADN59752	Peptide-v	786	31	91.2	412	5	AAE29083	Aae29083	Human tra
714	31	91.2	41	AAy96528	Thrombopo	787	31	91.2	412	5	AAU77103	Aau77103	Human tra
715	31	91.2	41	ABB73389	TPO-mimet	788	31	91.2	412	6	ADA66610	Ada66610	Human tra
716	31	91.2	41	ABB73388	TPO-mimet	789	31	91.2	412	6	ABU54472	Abu54472	Human tum
717	31	91.2	42	AAy96530	Thrombopo	790	31	91.2	412	8	ADH11592	Adh11592	Human bon
718	31	91.2	42	AAB17296	TPO-mimet	791	31	91.2	412	8	ADS87980	Ads87980	Tumour tr
719	31	91.2	42	AAB17308	Synthetic	792	31	91.2	412	8	ADU18086	Adul8086	Human can
720	31	91.2	42	AAB17282	TPO-mimet	793	31	91.2	412	9	ADY27744	Ady27744	Human tra
721	31	91.2	42	AAB17281	TPO-mimet	794	31	91.2	412	9	AEB01311	Aeb01311	Human TGF
722	31	91.2	42	ABB73404	TMP-TMP g	795	31	91.2	412	9	AEB44146	Aeb44146	Human tra
723	31	91.2	42	ADN59751	Peptide-v	796	31	91.2	413	2	AAR22038	Aar22038	Mutant Tr
724	31	91.2	43	ADN59752	Peptide-v	797	31	91.2	414	3	AAy32289	Aay32289	Pseudomon
725	31	91.2	43	ADN59759	Peptide-v	798	31	91.2	414	3	AAy93565	Aay93565	Amino aci
726	31	91.2	44	ADN59817	Peptide-	799	31	91.2	414	4	AAy72225	Aay72225	Pseudomon
727	31	91.2	60	AAB17311	Synthetic	800	31	91.2	414	4	AAy72237	Aay72237	Pseudomon
728	31	91.2	60	ABB73405	TMP-TMP g	801	31	91.2	414	4	AAy72245	Aay72245	Pseudomon
729	31	91.2	68	ABB39037	Peptide #	802	31	91.2	414	4	AAy72246	Aay72246	Pseudomon
730	31	91.2	68	AAM32526	Peptide #	803	31	91.2	414	4	AAy72265	Aay72265	Pseudomon
731	31	91.2	68	AAM72267	Human bon	804	31	91.2	414	4	AAy72272	Aay72272	Pseudomon
732	31	91.2	68	ABG53953	Human liv	805	31	91.2	414	4	AAy72279	Aay72279	Pseudomon
733	31	91.2	68	ABG42082	Human pep	806	31	91.2	414	4	AAy72291	Aay72291	Pseudomon
734	31	91.2	114	AAB25241	Eucalyptu	807	31	91.2	414	4	AAy72300	Aay72300	Pseudomon
735	31	91.2	122	ADV44474	Anti-teta	808	31	91.2	414	4	AAy72301	Aay72301	Pseudomon
736	31	91.2	122	AEB12946	Antibody	809	31	91.2	414	4	AAy72336	Aay72336	p. putida
737	31	91.2	128	ADQ16705	Modified	810	31	91.2	414	4	AAy72213	Aay72213	Pseudomon
738	31	91.2	128	ADV44466	Anti-teta	811	31	91.2	414	4	AAy72214	Aay72214	Pseudomon
739	31	91.2	128	ADV44463	Anti-teta	812	31	91.2	414	4	AAy72220	Aay72220	Pseudomon
740	31	91.2	128	ADV44467	Anti-teta	813	31	91.2	414	4	AAy72226	Aay72226	Pseudomon
741	31	91.2	128	ADV44465	Anti-teta	814	31	91.2	414	4	AAy72231	Aay72231	Pseudomon
742	31	91.2	128	ADV44441	pAX116 va	815	31	91.2	414	4	AAy72234	Aay72234	Pseudomon
743	31	91.2	128	ADV44464	Anti-teta	816	31	91.2	414	4	AAy72240	Aay72240	Pseudomon
744	31	91.2	128	AEB12934	Antibody	817	31	91.2	414	4	AAy72244	Aay72244	Pseudomon
745	31	91.2	128	AEB12935	Antibody	818	31	91.2	414	4	AAy72247	Aay72247	Pseudomon
746	31	91.2	128	AEB12913	Antibody	819	31	91.2	414	4	AAy72266	Aay72266	Pseudomon
747	31	91.2	128	AEB12936	Antibody	820	31	91.2	414	4	AAy72298	Aay72298	Pseudomon
748	31	91.2	128	AEB12937	Antibody	821	31	91.2	414	4	AAy72319	Aay72319	Pseudomon
749	31	91.2	128	AEB12938	Antibody	822	31	91.2	414	4	AAy72338	Aay72338	p. putida
750	31	91.2	129	ABG71751	Antibody	823	31	91.2	414	4	AAy72255	Aay72255	Pseudomon
751	31	91.2	131	ABG71753	Antibody	824	31	91.2	414	4	AAy72270	Aay72270	Pseudomon
752	31	91.2	132	ADV44473	Anti-teta	825	31	91.2	414	4	AAy72273	Aay72273	Pseudomon
753	31	91.2	133	ABG71752	Antibody	826	31	91.2	414	4	AAy72274	Aay72274	Pseudomon
754	31	91.2	135	ABG71749	Antibody	827	31	91.2	414	4	AAy72275	Aay72275	Pseudomon

828 31 91.2 414 4 AAY72306 Aay72306 Pseudomon 901
829 31 91.2 414 4 AAY72309 Aay72309 Pseudomon 902
830 31 91.2 414 4 AAY72312 Aay72312 Pseudomon 903
831 31 91.2 414 4 AAY72331 Aay72331 Pseudomon 904
832 31 91.2 414 4 AAY72233 Aay72233 Pseudomon 905
833 31 91.2 414 4 AAY72236 Aay72236 Pseudomon 906
834 31 91.2 414 4 AAY72260 Aay72260 Pseudomon 907
835 31 91.2 414 4 AAY72261 Aay72261 Pseudomon 908
836 31 91.2 414 4 AAY72276 Aay72276 Pseudomon 909
837 31 91.2 414 4 AAY72278 Aay72278 Pseudomon 910
838 31 91.2 414 4 AAY72305 Aay72305 Pseudomon 911
839 31 91.2 414 4 AAY72317 Aay72317 Pseudomon 912
840 31 91.2 414 4 AAY72324 Aay72324 Pseudomon 913
841 31 91.2 414 4 AAY72328 Aay72328 Pseudomon 914
842 31 91.2 414 4 AAY72232 Aay72232 Pseudomon 915
843 31 91.2 414 4 AAY72241 Aay72241 Pseudomon 916
844 31 91.2 414 4 AAY72252 Aay72252 Pseudomon 917
845 31 91.2 414 4 AAY72287 Aay72287 Pseudomon 918
846 31 91.2 414 4 AAY72311 Aay72311 Pseudomon 919
847 31 91.2 414 4 AAY72321 Aay72321 Pseudomon 920
848 31 91.2 414 4 AAY72207 Aay72207 Pseudomon 921
849 31 91.2 414 4 AAY72219 Aay72219 Pseudomon 922
850 31 91.2 414 4 AAY72221 Aay72221 Pseudomon 923
851 31 91.2 414 4 AAY72223 Aay72223 Pseudomon 924
852 31 91.2 414 4 AAY72254 Aay72254 Pseudomon 925
853 31 91.2 414 4 AAY72263 Aay72263 Pseudomon 926
854 31 91.2 414 4 AAY72281 Aay72281 Pseudomon 927
855 31 91.2 414 4 AAY72292 Aay72292 Pseudomon 928
856 31 91.2 414 4 AAY72310 Aay72310 Pseudomon 929
857 31 91.2 414 4 AAY72216 Aay72216 Pseudomon 930
858 31 91.2 414 4 AAY72227 Aay72227 Pseudomon 931
859 31 91.2 414 4 AAY72248 Aay72248 Pseudomon 932
860 31 91.2 414 4 AAY72282 Aay72282 Pseudomon 933
861 31 91.2 414 4 AAY72283 Aay72283 Pseudomon 934
862 31 91.2 414 4 AAY72326 Aay72326 Pseudomon 935
863 31 91.2 414 4 AAY72235 Aay72235 Pseudomon 936
864 31 91.2 414 4 AAY72249 Aay72249 Pseudomon 937
865 31 91.2 414 4 AAY72284 Aay72284 Pseudomon 938
866 31 91.2 414 4 AAY72308 Aay72308 Pseudomon 939
867 31 91.2 414 4 AAY72322 Aay72322 Pseudomon 940
868 31 91.2 414 4 AAY72474 Aay72474 Pseudomon 941
869 31 91.2 414 4 AAY72215 Aay72215 Pseudomon 942
870 31 91.2 414 4 AAY72238 Aay72238 Pseudomon 943
871 31 91.2 414 4 AAY72243 Aay72243 Pseudomon 944
872 31 91.2 414 4 AAY72251 Aay72251 Pseudomon 945
873 31 91.2 414 4 AAY72264 Aay72264 Pseudomon 946
874 31 91.2 414 4 AAY72293 Aay72293 Pseudomon 947
875 31 91.2 414 4 AAY72293 Aay72293 Pseudomon 948
876 31 91.2 414 4 AAY72323 Aay72323 Pseudomon 949
877 31 91.2 414 4 AAY72256 Aay72256 Pseudomon 950
878 31 91.2 414 4 AAY72258 Aay72258 Pseudomon 951
879 31 91.2 414 4 AAY72262 Aay72262 Pseudomon 952
880 31 91.2 414 4 AAY72271 Aay72271 Pseudomon 953
881 31 91.2 414 4 AAY72277 Aay72277 Pseudomon 954
882 31 91.2 414 4 AAY72334 Aay72334 Pseudomon 955
883 31 91.2 414 4 AAY72218 Aay72218 Pseudomon 956
884 31 91.2 414 4 AAY72222 Aay72222 Pseudomon 957
885 31 91.2 414 4 AAY72228 Aay72228 Pseudomon 958
886 31 91.2 414 4 AAY72230 Aay72230 Pseudomon 959
887 31 91.2 414 4 AAY72257 Aay72257 Pseudomon 960
888 31 91.2 414 4 AAY72304 Aay72304 Pseudomon 961
889 31 91.2 414 4 AAY72333 Aay72333 Pseudomon 962
890 31 91.2 414 4 AAY72267 Aay72267 Pseudomon 963
891 31 91.2 414 4 AAY72286 Aay72286 Pseudomon 964
892 31 91.2 414 4 AAY72288 Aay72288 Pseudomon 965
893 31 91.2 414 4 AAY72290 Aay72290 Pseudomon 966
894 31 91.2 414 4 AAY72313 Aay72313 Pseudomon 967
895 31 91.2 414 4 AAY72315 Aay72315 Pseudomon 968
896 31 91.2 414 4 AAY72316 Aay72316 Pseudomon 969
897 31 91.2 414 4 AAY72325 Aay72325 Pseudomon 970
898 31 91.2 414 4 AAY72212 Aay72212 Pseudomon 971
899 31 91.2 414 4 AAY72259 Aay72259 Pseudomon 972
900 31 91.2 414 4 AAY72294 Aay72294 Pseudomon 973
Aay72297 Pseudomon

31 91.2 414 4 AAY72314 Aay72314 Pseudomon
31 91.2 414 4 AAY72318 Aay72318 Pseudomon
31 91.2 414 4 AAY72327 Aay72327 Pseudomon
31 91.2 414 4 AAY72229 Aay72229 Pseudomon
31 91.2 414 4 AAY72242 Aay72242 Pseudomon
31 91.2 414 4 AAY72250 Aay72250 Pseudomon
31 91.2 414 4 AAY72285 Aay72285 Pseudomon
31 91.2 414 4 AAY72289 Aay72289 Pseudomon
31 91.2 414 4 AAY72295 Aay72295 Pseudomon
31 91.2 414 4 AAY72299 Aay72299 Pseudomon
31 91.2 414 4 AAY72320 Aay72320 Pseudomon
31 91.2 414 4 AAY72335 Aay72335 Pseudomon
31 91.2 414 4 AAY72337 P. putida
31 91.2 414 4 AAY72217 Pseudomon
31 91.2 414 4 AAY72253 Pseudomon
31 91.2 414 4 AAY72269 Pseudomon
31 91.2 414 4 AAY72303 Pseudomon
31 91.2 414 4 AAY72307 Pseudomon
31 91.2 414 4 AAY72224 Pseudomon
31 91.2 414 4 AAY72239 Pseudomon
31 91.2 414 4 AAY72268 Pseudomon
31 91.2 414 4 AAY72280 Pseudomon
31 91.2 414 4 AAY72296 Pseudomon
31 91.2 414 4 AAY72302 Pseudomon
31 91.2 414 4 AAY72329 Pseudomon
31 91.2 414 4 AAY72330 Pseudomon
31 91.2 414 4 AAY72332 Pseudomon
31 91.2 414 5 ABB08123 P. putida
31 91.2 414 5 ABB08127 P. putida
31 91.2 414 5 ABB08126 P. putida
31 91.2 414 5 ABB08128 P. putida
31 91.2 414 6 ABR43996 P. putida
31 91.2 414 8 ADO47232 Protease-
31 91.2 414 9 ADV90218 P. putida
31 91.2 414 9 ADV69262 Murine TG
31 91.2 423 8 ADQ17024 Murine TG
31 91.2 427 8 ADQ17026 Murine TG
31 91.2 456 2 AAR46228 Pig TGF-b
31 91.2 456 2 AAW78786 Pig trans
31 91.2 457 2 AAR04080 Polypepti
31 91.2 459 9 ADV44459 Anti-teta
31 91.2 472 5 ABP51695 5G1.1-TPO
31 91.2 472 8 ADQ16647 Immunoglo
31 91.2 472 9 ADV44385 5G1.1 hea
31 91.2 472 9 AEB12858 Antibody
31 91.2 531 6 ABU29925 Protein e
31 91.2 543 7 ADC94090 E. faeciu
31 91.2 602 7 ADE59823 Human Pro
31 91.2 637 7 ABO66004 Klebsiell
31 91.2 653 8 ADS30548 Bacterial
31 91.2 744 9 AEB40488 L. pneumo
31 91.2 751 9 AEB37147 L. pneumo
31 91.2 984 5 ABB93326 Herbicida
31 91.2 1043 6 ABU15501 Protein e
31 91.2 1094 4 AAG81238 Mycobacte
31 91.2 1151 7 ABO74850 Pseudomon
31 91.2 1220 6 ABU41578 Protein e
31 91.2 1630 4 ABB44550 Human wou
31 91.2 1632 6 ABP98331 Amino aci
31 91.2 2171 7 ADC99058 Human pol
31 91.2 2955 4 ABG21509 Novel KPP
31 91.2 2988 4 ABB44617 Human hum
31 91.2 4342 4 AAU33611 Pseudomon
31 91.2 4342 6 ABU15500 Protein e
31 88.2 12 2 AAW09582 Thrombopo
31 88.2 12 2 AAW36733 Thrombopo
31 88.2 12 4 AAU25952 Human thr
31 88.2 16 2 AAW09609 Thrombopo
31 88.2 16 2 AAW36760 Thrombopo
31 88.2 16 4 AAU25979 Human thr
31 88.2 18 2 AAW00367 Streptomy
31 88.2 19 2 AAW09495 Thrombopo

974 30 88.2 19 2 AAW36646 Aaw36646 Thrombopo
975 30 88.2 19 3 AAB17023 Aab17023 TPO-mimet
976 30 88.2 19 4 AAU25865 Aau25865 Human thr
977 30 88.2 19 5 ABB72909 Abb72909 TPO mimet
978 30 88.2 19 7 ADJ73061 Adj73061 TPO mimet
979 30 88.2 19 8 ADJ52696 Adj52696 CH1 delet
980 30 88.2 19 8 ADJ51657 Adj51657 CH1 delet
981 30 88.2 20 2 AAW06885 Aaw06885 Glycosami
982 30 88.2 20 2 AAR92270 Aar92270 Protein C
983 30 88.2 20 6 ABJ37568 Abj37568 Heparin b
984 30 88.2 20 9 AEB12129 Aeb12129 Antimicro
985 30 88.2 51 4 AAU54014 Aau54014 Propionib
986 30 88.2 51 6 ABM50533 Abm50533 Propionib
987 30 88.2 89 4 ABB33444 Abb33444 Peptide #
988 30 88.2 89 4 ABG48291 Abg48291 Human liv
989 30 88.2 117 8 ADQ67108 Adq67108 Novel hum
990 30 88.2 120 4 AAB68965 Aab68965 Adenoviru
991 30 88.2 123 4 AAU19758 Aau19758 Human nov
992 30 88.2 123 4 AAU21495 Aau21495 Human nov
993 30 88.2 123 5 ABP47978 Abp47978 Human pol
994 30 88.2 123 7 ADC10940 Adc10940 Human ext
995 30 88.2 136 4 ABG12413 Abg12413 Novel hum
996 30 88.2 163 3 AAG333763 Aag333763 Arabidops
997 30 88.2 166 3 AAG333762 Aag333762 Arabidops
998 30 88.2 177 3 AAG333761 Aag333761 Arabidops
999 30 88.2 178 3 AAG54807 Aag54807 Arabidops
1000 30 88.2 181 6 ABM71724 Abm71724 Staphyloc

ALIGNMENTS

RESULT 1
AAB16978
ID AAB16978 standard; peptide; 6 AA.

XX AAB16978;
DT 31-OCT-2000 (first entry)
XX TPO-mimetic peptide sequence SEQ ID NO:34.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS WO200024782-A2.
XX 04-MAY-2000.
PD 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX (AMGE-) AMGEN INC.
PA Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
DR Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
PT Claim 19, Page 207; 608pp; English.
XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX Sequence 6 AA;

Query Match 100.0%; Score 34; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1 TLREWL 6

RESULT 2
ABB72864
ID ABB72864 standard; peptide; 6 AA.

XX ABB72864;
AC ABB72864;
XX 05-APR-2002 (first entry)
XX TPO mimetic peptide SEQ ID NO:34.

DE Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX Homo sapiens.
OS Synthetic.
XX WO200183525-A2.
PN 08-NOV-2001.
PD 02-MAY-2001; 2001WO-US014310.
XX 03-MAY-2000; 2000US-00563286.
PR (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
PT

PS	Claim 39; Page 43; 176pp; English.	PT	preventing hematopoietic or immune disorders, e.g. thrombocytopenia.
XX		XX	Disclosure; Page 11; 97pp; English.
CC	The present invention describes a vehicle-peptide molecule (I) or its	CC	The invention discloses a variant of an immunoglobulin (Ig) variable
CC	multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,	CC	heavy or light chain domain that comprises at least one complementarity
CC	cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,	CC	determining region (CDR) and framework regions flanking the CDR. The CDR
CC	antianaemic, anorectic, antiinfertility, haemostatic, dermatological and	CC	also has added or substituted to it, at least one binding sequence which
CC	neuroprotective activities. (I) can be used as a therapeutic or	CC	is heterologous to the CDR and is an antigenic, agonistic sequence from a
CC	prophylactic agent as well as for screening purposes. (I) is useful for	CC	thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
CC	diagnosing diseases characterised by dysfunction of their associated	CC	can be a binding sequence heterologous to the CDR, a cytotoxic T-
CC	protein of interest, for identifying normal or abnormal proteins of	CC	lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
CC	interest, as a part of diagnostic kit to detect the presence of their	CC	cell sequence or a combination of each. The variant or thrombopoietin
CC	proteins of interest in a biological sample. Additionally, (I) is useful	CC	synthetbody, pharmaceutical and vaccine compositions are useful for
CC	for treating inflammatory and autoimmune diseases, tumour growth, cancer,	CC	stimulating proliferation, growth or differentiation of haematopoietic
CC	rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,	CC	cells, particularly platelet progenitor cells. The variants are also
CC	infertility, and neurological degenerative diseases. (I), comprising EPO-	CC	useful for treating or preventing haematopoietic or immune disorders
CC	mimetic compounds are useful for treating disorders characterised by low	CC	resulting from chemotherapy, radiation therapy, or bone marrow
CC	red blood cell levels such as anaemia. The TPO-mimetic comprising	CC	transfusions (e.g. thrombocytopenia or disseminated intravascular
CC	compounds are useful for treating conditions that involve an existing	CC	coagulation). Compositions comprising the synthetbodies can be used for
CC	megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet	CC	the mobilisation, amplification and ex vivo expansion of stem cells and
CC	deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic	CC	committed precursor cells for autologous and allogeneic transplantation
CC	tumour which result in thrombocytopaenia, systemic lupus erythematosus,	CC	as well as for the expansion of stem cells for gene therapy. They are
CC	and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777	CC	also useful as diagnostic or analytical reagents for studying the
CC	represent amino acid and nucleic acid sequences used in the	CC	function of thrombopoietin and its receptor in vivo or in vitro. The
CC	exemplification of the present invention	CC	sequence presented is the TPO receptor (MPL) agonist peptide consensus
XX		CC	sequence #2
SQ	Sequence 6 AA;	XX	
	Query Match 100.0%; Score 34; DB 5; Length 6;	SQ	Sequence 6 AA;
	Best Local Similarity 100.0%; Pred. No. 2e+06;		Query Match 100.0%; Score 34; DB 6; Length 6;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Best Local Similarity 100.0%; Pred. No. 2e+06;
			Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 TLREWL 6	Qy	1 TLREWL 6
Db	1 TLREWL 6	Db	1 TLREWL 6
RESULT 3		RESULT 4	
ABG711746		ADJ73015	
ID	ABG711746 standard; peptide; 6 AA.	ID	ADJ73015 standard; peptide; 6 AA.
XX		XX	
AC	ABG711746;	AC	ADJ73015;
XX		XX	
DT	20-JAN-2003 (first entry)	DT	06-MAY-2004 (first entry)
XX		XX	
DE	TPO receptor, MPL, agonist peptide consensus sequence #2.	XX	TPO mimetic peptide sequence SeqID 469.
KW	Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;	XX	mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW	complementarity determining region; CDR; antigenic; thrombopoietin; TPO;	XX	cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW	thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;	KW	immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW	T-helper cell; B-helper cell; synthetbody; pharmaceutical; vaccine;	KW	TPO.
KW	proliferation; growth; differentiation; haematopoietic cell;	XX	
KW	platelet progenitor cell; immune disorder; thrombocytopenia;	OS	Synthetic.
KW	disseminated intravascular coagulation; stem cell; transplantation;	XX	
KW	gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.	XX	WO2003084477-A2.
OS		PN	
XX	Synthetic.	XX	16-OCT-2003.
XX		PD	
PN	WO200278612-A2.	XX	24-MAR-2003; 2003WO-US009139.
XX		PF	
PD	10-OCT-2002.	XX	29-MAR-2002; 2002US-0368791P.
XX		PR	
PF	02-APR-2002; 2002WO-US010301.	XX	(CENZ) CENTOCOR INC.
XX		PA	
XX	02-APR-2001; 2001US-0281183P.	XX	Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX	(PURD) PURDUE PHARMA LP.	XX	WPI; 2003-804237/75.
PA		XX	
XX	Soltis DA, Burch RM, Ogert RA;	XX	
PI		DR	
XX	WPI; 2003-040615/03.	XX	
DR		PT	New CDR mimetibody comprising a portion of a heavy or light chain
XX		PT	variable region comprising human framework or ligand binding region,
PT	New thrombopoietin synthetbodies, useful for stimulating proliferation,	PT	useful for preparing a composition for treating e.g., immune,
PT	growth, or differentiation of hematopoietic cells, for treating or		

PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 469; 97pp; English.

PS This invention relates to novel mammalian CDR mimetibodies, specific

XX portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is a TPO mimetic peptide sequence used to make a

CC mimetibody of the invention.

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 7; Length 6;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db |||||

1 TLREWL 6

RESULT 5

ADJ52650

ID ADJ52650 standard; peptide; 6 AA.

XX

AC ADJ52650;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID469.

XX

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082870/08.

XX

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX

PS Claim 2; SEQ ID NO 469; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an immunosuppressive,

CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,

CC antibacterial, virucide or fungicide activity. In addition, the disclosed

CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody

CC is useful for diagnosing or treating a disease condition in a cell,

CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart

CC failure), or neurodegenerative (for example multiple sclerosis, dementia

CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or

CC fungal infection). The present sequence is that of a peptide which may be

CC used during the creation of a mimetibody of the invention.

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 8; Length 6;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db |||||

1 TLREWL 6

RESULT 6

ADJ51611

ID ADJ51611 standard; peptide; 6 AA.

XX

AC ADJ51611;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID469.

XX

KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunological disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002424-A2.

XX

PD 08-JAN-2004.

XX

PF 30-JUN-2003; 2003WO-US020495.

XX

PR 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082872/08.

XX

PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic, and
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 14; SEQ ID NO 469; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1 TLREWL 6

RESULT 7
ADJ73040
ID ADJ73040 standard; peptide; 9 AA.
XX
XX
AC ADJ73040;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 494.
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
XX (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX

New CDR mimetibody comprising a portion of a heavy or light chain
variable region comprising human framework or ligand binding region,
useful for preparing a composition for treating e.g., immune,
cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 494; 97pp; English.
PS
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 8
ADJ73037
ID ADJ73037 standard; peptide; 9 AA.
XX
XX
AC ADJ73037;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 491.
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
XX (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX

New CDR mimetibody comprising a portion of a heavy or light chain
variable region comprising human framework or ligand binding region,
useful for preparing a composition for treating e.g., immune,
cardiovascular or neurologic disease.

Disclosure; SEQ ID NO 491; 97pp; English.
XX
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 9 AA;
SQ

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 9
ADJ73041
ID ADJ73041 standard; peptide; 9 AA.
XX
AC ADJ73041;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 495.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 495; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 9 AA;
SQ

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 10
ADJ73039
ID ADJ73039 standard; peptide; 9 AA.
XX
AC ADJ73039;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 493.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 493; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 9 AA;
SQ

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 11
ADJ73038
ID ADJ73038 standard; peptide; 9 AA.

XX AC ADJ73038;
XX 06-MAY-2004 (first entry)
XX TPO mimetic peptide sequence SeqID 492.
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX OS Synthetic.

XX PN WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX DR WPI; 2003-804237/75.

XX PT New CDR mimetibody comprising a portion of a heavy or light chain
variable region comprising human framework or ligand binding region,
useful for preparing a composition for treating e.g., immune,
cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 492; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific
portions or variants thereof. Specifically, it refers to an antibody
fragment where a protein has been inserted into, or replaces a portion
of, one or more CDR regions, such that each CDR mimetibody comprises at
least one portion of a heavy chain or light chain variable region, which
itself comprises at least one human framework region and at least one
ligand binding region (LBR). The present invention describes human
mimetibodies, including modified immunoglobulins and cleavage products
that can be useful in gene therapy and the generation of transgenic
plants and animals. Furthermore, the CDR mimetibody is useful for
preparing compositions for modulating, treating or reducing the symptoms
of immune, cardiovascular, infectious, malignant and/ or neurologic
diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
cardiant, antimicrobial, cytostatic and neuroprotective activities. This
peptide sequence is a TPO mimetic peptide sequence used to make a
mimetibody of the invention.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||

Db 2 TLREWL 7
RESULT 12
AAW09471
ID AAW09471 standard; protein; 10 AA.
XX AC AAW09471;
XX 10-SEP-1997 (first entry)
XX Thrombopoietin receptor binding peptide.
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.

XX OS Synthetic.

XX PN WO9640189-A1.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.

XX PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
mimetic(s) - useful in treatment of haematological disorders, esp.
thrombocytopenia resulting from chemotherapy, etc.

XX PS Disclosure; Page 26; 106pp; English.

XX CC The present sequence is a peptide which binds to thrombopoietin (TPO)
receptor (TR). The compound can be used for treating patients suffering
from haematological disorders and thrombocytopenia resulting from
chemotherapy, radiation therapy or bone marrow transfusions. The peptide
may also be used to maintain the proliferation and growth of TPO-
dependent cell lines and for use in biological research, for detecting
TPO receptors on living cells

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||

Db 5 TLREWL 10

RESULT 13
AAW09566

ID AAW09566 standard; protein; 10 AA.

XX AC AAW09566;

XX DT 10-SEP-1997 (first entry)

XX DE Thrombopoietin receptor binding peptide.

KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7

RESULT 14
AAW36622
ID AAW36622 standard; peptide; 10 AA.
XX
AC AAW36622;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 5 TLREWL 10

RESULT 15
AAW36717
ID AAW36717 standard; peptide; 10 AA.
XX
AC AAW36717;
XX
DT 27-FEB-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX

PS Disclosure; Page 29; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 16
AAB17000
ID AAB17000 standard; peptide; 10 AA.
XX
AC AAB17000;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:56.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.

XX Synthetic.
OS
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 214; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX Sequence 10 AA;

Query Match 100.0%; Score 34; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 17
AAU25841
ID AAU25841 standard; peptide; 10 AA.

XX
AC AAU25841;

XX
DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #27.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.

PS Disclosure; Col 19; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
| | | | |
Db 5 TLREWL 10

RESULT 18
AAU25936
ID AAU25936 standard; peptide; 10 AA.
XX
AC AAU25936;

XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #122.
DE
XX

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX

OS Homo sapiens.

XX
PN U96251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 22; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent hematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
| | | | |
Db 2 TLREWL 7

RESULT 19
ABB72886
ID ABB72886 standard; peptide; 10 AA.
XX
AC ABB72886;

XX 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:56.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 20
ADJ52672

ID ADJ52672 standard; peptide; 10 AA.

XX
AC ADJ52672;

XX
DT 06-MAY-2004 (first entry)

XX
DE CH1 deleted mimetibody-related peptide SeqID491.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.
OS Synthetic.

XX Key Location/Qualifiers

FH Misc-difference 1. .10
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX WO2004002417-A2.

XX
PN 08-JAN-2004.

XX
PD 27-JUN-2003; 2003WO-US020347.

XX
PF 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;

XX WPI; 2004-082870/08.

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 491; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 21
ADJ51633

ID ADJ51633 standard; peptide; 10 AA.

XX
AC ADJ51633;

XX
DT 06-MAY-2004 (first entry)

XX
DE CH1 deleted mimetibody-related peptide SeqID491.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX Unidentified.
OS Synthetic.

XX Key Location/Qualifiers

FH Misc-difference 1. .10
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX WO2004002424-A2.

XX
PD 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

XX 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavyner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX WPI; 2004-082872/08.
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX Claim 14; SEQ ID NO 491; 123pp; English.
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 22
AAB17001
ID AAB17001 standard; peptide; 11 AA.
XX AAB17001;
AC AAB17001;

XX 31-OCT-2000 (first entry)
XX TPO-mimetic peptide sequence SEQ ID NO:57.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.

XX OS Synthetic.
OS WO200024782-A2.
XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.

PA (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX Claim 19; Page 214; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 34; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 23
ABB72887
ID ABB72887 standard; peptide; 11 AA.
XX ABB72887;
AC ABB72887;

XX 05-APR-2002 (first entry)
XX TPO mimetic peptide SEQ ID NO:57.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX OS Homo sapiens.
OS Synthetic.

XX WO200183525-A2.
XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

PR 03-MAY-2000; 2000US-00563286.
XX (AMGE-) AMGEN INC.
PA
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;
Query Match 100.0%; Score 34; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db |||||
2 TLREWL 7
RESULT 24
ADJ52673
ID ADJ52673 standard; peptide; 11 AA.
XX
AC ADJ52673;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID492.
XX
KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..11 /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX WO2004002417-A2.
PN
XX 08-JAN-2004.
PD
XX 27-JUN-2003; 2003WO-US020347.
PF
XX 28-JUN-2002; 2002US-0392431P.
PR (CENZ) CENTOCOR INC.
XX
PA Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
XX Kutoloski KA;
PI
XX WPI; 2004-082870/08.
DR
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
PT
XX Claim 2; SEQ ID NO 492; 129pp; English.
PS
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 11 AA;
Query Match 100.0%; Score 34; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db |||||
2 TLREWL 7
RESULT 25
ADJ51634
ID ADJ51634 standard; peptide; 11 AA.
XX
AC ADJ51634;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID492.
XX
KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 1. .11
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
PA
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloeki KA;
PI
XX WPI; 2004-082872/08.
DR
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 492; 123pp; English.
XX
CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 11 AA;
Query Match 100.0%; Score 34; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB 2 TLREWL 7
RESULT 26
AAW09476
ID AAW09476 standard; protein; 12 AA.
XX
AC AAW09476;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.

XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB 5 TLREWL 10
RESULT 27
AAW09581
ID AAW09581 standard; protein; 12 AA.
XX
AC AAW09581;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 12 AA;
XX
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 28
AAW09577
ID AAW09577 standard; protein; 12 AA.
XX
AC AAW09577;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX

PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 12 AA;
XX
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 29
AAW09524
ID AAW09524 standard; protein; 12 AA.
XX
AC AAW09524;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 28; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells

XX WO9640750-A1.
PN 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009623.
PD 07-JUN-1995; 95US-00478128.
PF 07-JUN-1995; 95US-00485301.
XX (GLAX) GLAXO GROUP LTD.
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
DR Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
PT
XX Disclosure; Page 28; 106pp; English.
PS The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 33
AAB17002
ID AAB17002 standard; peptide; 12 AA.
XX AAB17002;
AC
XX 31-OCT-2000 (first entry)
DT
XX TPO-mimetic peptide sequence SEQ ID NO:58.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS
XX WO200024782-A2.
PN
XX 04-MAY-2000.
PD
XX 25-OCT-1999; 99WO-US025044.
PF
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.
PR
XX (AMGE-) AMGEN INC.
PA

XX Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically
DR active peptides, useful for treating cancer and autoimmune diseases.
XX Claim 19; Page 214; 608pp; English.
PT The present invention describes composition of matter (I) comprising an
PT Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX P3, and P4 = are each independently sequences of pharmacologically active
XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX cells from the present invention can be used for producing pharmaceutical
XX compositions. The compositions are useful for treating cancer, asthma,
XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX a Fab domain) can provide a longer half-life or incorporate functions
XX such as Fc receptor binding, protein A binding, complement fixation, and
XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX AAB18003 represent nucleotide and amino acid sequences used in the
XX exemplification of the present invention
SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 34
AAB16993
ID AAB16993 standard; peptide; 12 AA.
XX AAB16993;
AC
XX 31-OCT-2000 (first entry)
DT
XX TPO-mimetic peptide sequence SEQ ID NO:49.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS
XX WO200024782-A2.
PN
XX 04-MAY-2000.
PD
XX 25-OCT-1999; 99WO-US025044.
PF
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.
PR
XX (AMGE-) AMGEN INC.
PA Feige U, Liu C, Cheetham J, Boone TC;
XX

DR WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.

XX

PS Claim 19; Page 212; 608pp; English.

XX

CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each

CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active

CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic,

CC thrombolytic and immunosuppressive activities. DNAs, vectors and host

CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,

CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to

CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
Db 2 TLREWL 7

RESULT 35
AAU25846
ID AAU25846 standard; peptide; 12 AA.

XX AAU25846;

XX 17-DEC-2001 (first entry)

DT Human thrombopoietin receptor (TPO-R) activator peptide #32.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

OS US6251864-B1.

PN 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI Yin Q;

DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX

PS Disclosure; Col 19; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent haematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

CC progenitors alone or in conjunction with additional cytokines

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
Db 5 TLREWL 10

RESULT 36
AAU25951

ID AAU25951 standard; peptide; 12 AA.

XX AAU25951;

XX 17-DEC-2001 (first entry)

DT Human thrombopoietin receptor (TPO-R) activator peptide #137.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

OS US6251864-B1.

PN 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI Yin Q;

XX WPI; 2001-564142/63.
DR
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 22; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk.The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 37
AAU25894
ID AAU25894 standard; peptide; 12 AA.
XX
AC AAU25894;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #80.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 21; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk.The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 38
AAU25947
ID AAU25947 standard; peptide; 12 AA.
XX
AC AAU25947;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #133.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
DR
XX
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX
PS Disclosure; Col 22; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked, immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 39
ABB72888
ID ABB72888 standard; peptide; 12 AA.

XX ABB72888;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:58.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX

PR 03-MAY-2000; 2000US-00563286.
XX (AMGE-) AMGEN INC.
PA
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
PI WPI; 2002-130313/17.
XX
DR Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
PT
XX Claim 39; Page 43; 176pp; English.
PS
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytosstatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 40

ABB72879
ID ABB72879 standard; peptide; 12 AA.

XX ABB72879;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:49.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS Synthetic.

XX PN WO200183525-A2.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-US014310.
XX PR 03-MAY-2000; 2000US-00563286.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX DR WPI; 2002-130313/17.
XX PT Novel vehicle-peptide molecule or its multimers useful for treating
XX PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX PS Claim 39; Page 43; 176pp; English.
XX CC The present invention describes a vehicle-peptide molecule (I) or its
XX CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
XX CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
XX CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
XX CC neuroprotective activities. (I) can be used as a therapeutic or
XX CC prophylactic agent as well as for screening purposes. (I) is useful for
XX CC diagnosing diseases characterised by dysfunction of their associated
XX CC protein of interest, for identifying normal or abnormal proteins of
XX CC interest, as a part of diagnostic kit to detect the presence of their
XX CC proteins of interest in a biological sample. Additionally, (I) is useful
XX CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
XX CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
XX CC infertility, and neurological degenerative diseases. (I), comprising EPO-
XX CC mimetic compounds are useful for treating disorders characterised by low
XX CC red blood cell levels such as anaemia. The TPO-mimetic comprising
XX CC compounds are useful for treating conditions that involve an existing
XX CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
XX CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
XX CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
XX CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
XX CC represent amino acid and nucleic acid sequences used in the
XX CC exemplification of the present invention
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB |||||
2 TLREWL 7
RESULT 41
ADJ73030
ID ADJ73030 standard; peptide; 12 AA.
XX AC ADJ73030;
XX DT 06-MAY-2004 (first entry)
XX DE TPO mimetic peptide sequence SeqID 484.
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
XX TPO.
XX OS Synthetic.
XX PN WO2003084477-A2.

XX PD WO200183525-A2.
XX PF 08-NOV-2001.
XX PR 02-MAY-2001; 2001WO-US014310.
XX PR 03-MAY-2000; 2000US-00563286.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX DR WPI; 2002-130313/17.
XX PT Novel vehicle-peptide molecule or its multimers useful for treating
XX PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX PS Claim 39; Page 43; 176pp; English.
XX CC The present invention describes a vehicle-peptide molecule (I) or its
XX CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
XX CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
XX CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
XX CC neuroprotective activities. (I) can be used as a therapeutic or
XX CC prophylactic agent as well as for screening purposes. (I) is useful for
XX CC diagnosing diseases characterised by dysfunction of their associated
XX CC protein of interest, for identifying normal or abnormal proteins of
XX CC interest, as a part of diagnostic kit to detect the presence of their
XX CC proteins of interest in a biological sample. Additionally, (I) is useful
XX CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
XX CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
XX CC infertility, and neurological degenerative diseases. (I), comprising EPO-
XX CC mimetic compounds are useful for treating disorders characterised by low
XX CC red blood cell levels such as anaemia. The TPO-mimetic comprising
XX CC compounds are useful for treating conditions that involve an existing
XX CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
XX CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
XX CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
XX CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
XX CC represent amino acid and nucleic acid sequences used in the
XX CC exemplification of the present invention
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB |||||
2 TLREWL 7
RESULT 41
ADJ73030
ID ADJ73030 standard; peptide; 12 AA.
XX AC ADJ73030;
XX DT 06-MAY-2004 (first entry)
XX DE TPO mimetic peptide sequence SeqID 484.
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
XX TPO.
XX OS Synthetic.
XX PN WO2003084477-A2.

XX PD 16-OCT-2003.
XX PF 24-MAR-2003; 2003WO-US009139.
XX PR 29-MAR-2002; 2002US-0368791P.
XX PA (CENZ) CENTOCOR INC.
XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX DR WPI; 2003-804237/75.
XX PT New CDR mimetibody comprising a portion of a heavy or light chain
XX PT variable region comprising human framework or ligand binding region,
XX PT useful for preparing a composition for treating e.g., immune,
XX PT cardiovascular or neurologic disease.
XX PS Disclosure; SEQ ID NO 484; 97pp; English.
XX CC This invention relates to novel mammalian CDR mimetibodies, specific
XX CC portions or variants thereof. Specifically, it refers to an antibody
XX CC fragment where a protein has been inserted into, or replaces a portion
XX CC of, one or more CDR regions, such that each CDR mimetibody comprises at
XX CC least one portion of a heavy chain or light chain variable region, which
XX CC itself comprises at least one human framework region and at least one
XX CC ligand binding region (LBR). The present invention describes human
XX CC mimetibodies, including modified immunoglobulins and cleavage products
XX CC that can be useful in gene therapy and the generation of transgenic
XX CC plants and animals. Furthermore, the CDR mimetibody is useful for
XX CC preparing compositions for modulating, treating or reducing the symptoms
XX CC of immune, cardiovascular, infectious, malignant and/ or neurologic
XX CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
XX CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
XX CC peptide sequence is a TPO mimetic peptide sequence used to make a
XX CC mimetibody of the invention.
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB |||||
2 TLREWL 7
RESULT 42
ADJ52665
ID ADJ52665 standard; peptide; 12 AA.
XX AC ADJ52665;
XX DT 06-MAY-2004 (first entry)
XX DE CH1 deleted mimetibody-related peptide SeqID484.
XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
XX KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
XX KW fungicide; gene therapy; immune disorder; cardiovascular disease;
XX KW arrhythmia; hypertension; heart failure; neurodegenerative;
XX KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
XX KW cancerous condition; infectious disease; bacterial infection;
XX KW viral infection; fungal infection.
XX OS Unidentified.
XX OS Synthetic.
XX PN WO2004002417-A2.
XX PD 08-JAN-2004.
XX

PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 484; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 43
ADJ52674
ID ADJ52674 standard; peptide; 12 AA.
XX
AC ADJ52674;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID493.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX

PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 493; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db |||||
2 TLREWL 7

RESULT 44
ADJ51626
ID ADJ51626 standard; peptide; 12 AA.
XX
AC ADJ51626;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID484.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.

OS Synthetic.
XX WO2004002424-A2.
PN
XX
PD 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
PF
XX
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
PA
XX Heavner GA, Knight DM, Chrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
DR
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 484; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytosstatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 45
ADJ51635
ID ADJ51635 standard; peptide; 12 AA.
XX
AC ADJ51635;
XX
DT 06-MAY-2004 (first entry)
XX
XX CH1 deleted mimetibody-related peptide SeqID493.
DE
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytosstatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .12
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Chrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
XX
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 493; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytosstatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 46
AAW36732
ID AAW36732 standard; peptide; 13 AA.

XX AAW36732;
AC
XX
XX DT 09-MAR-1998 (first entry)
XX
XX DE Thrombopoietin receptor binding peptide.
XX
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
XX OS Synthetic.
XX
XX PN WO9640750-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 07-JUN-1996; 96WO-US009623.
XX
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX
XX PA (GLAXO) GLAXO GROUP LTD.
XX
XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX DR WPI; 1997-0522226/05.
XX
XX PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX PS Disclosure; Page 29; 106pp; English.
XX
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX SQ Sequence 13 AA;

Query Match 100.0%; Score 34; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 3 TLREWL 8

RESULT 47
AAB17015
ID AAB17015 standard; peptide; 13 AA.
XX
XX AC AAB17015;
XX
XX DT 31-OCT-2000 (first entry)
XX
XX DE TPO-mimetic peptide sequence SEQ ID NO:71.
XX
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.

XX OS Synthetic.
XX
XX PN WO200024782-A2.
XX
XX PD 04-MAY-2000.
XX
XX PF 25-OCT-1999; 99WO-US025044.
XX
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX
XX PA (AMGE-) AMGEN INC.
XX
XX PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX DR WPI; 2000-350702/30.
XX
XX PT Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
XX PS Claim 19; Page 219; 608pp; English.
XX
XX CC The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX P3, and P4 = are each independently sequences of pharmacologically active
XX peptides; L1, L2, L3, and L4 = are each independently 0 or 1, provided that at least 1
XX of a, d, e, and f = are each independently 0 or 1, provided that at least 1
XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX cells from the present invention can be used for producing pharmaceutical
XX compositions. The compositions are useful for treating cancer, asthma,
XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX a Fab domain) can provide a longer half-life or incorporate functions
XX such as Fc receptor binding, protein A binding, complement fixation, and
XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX AAB18003 represent nucleotide and amino acid sequences used in the
XX exemplification of the present invention
XX
XX SQ Sequence 13 AA;

Query Match 100.0%; Score 34; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 6 TLREWL 11

RESULT 48
AAB17003
ID AAB17003 standard; peptide; 13 AA.
XX
XX AC AAB17003;
XX
XX DT 31-OCT-2000 (first entry)
XX
XX DE TPO-mimetic peptide sequence SEQ ID NO:59.
XX
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
XX OS Synthetic.
XX

PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 215; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 34; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db 2 TLREWL 7
RESULT 49
ABB72901
ID ABB72901 standard; peptide; 13 AA.
XX
AC ABB72901;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:71.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 44; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 34; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db 6 TLREWL 11
RESULT 50
ABB72889
ID ABB72889 standard; peptide; 13 AA.
XX
AC ABB72889;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:59.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-US014310.
PF
XX
XX 03-MAY-2000; 2000US-00563286.
PR
XX
PA (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
PI
XX
XX WPI; 2002-130313/17.
DR
XX
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
PT
XX
PS Claim 39; Page 43; 176pp; English.
XX
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 34; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 2 TLREWL 7

Search completed: May 12, 2006, 10:36:42
Job time : 49.6897 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 7.28205 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-34
Perfect score: 34
Sequence: 1 TLREWL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	100.0	486	2	B86411	protein F3M18.4 [i
2	34	100.0	1045	2	S23570	pol polyprotein ho
3	32	94.1	473	2	E84853	hypothetical prote
4	32	94.1	560	2	S41808	glucose-6-phosphat
5	32	94.1	566	2	T09154	glucose-6-phosphat
6	32	94.1	568	2	S57830	glucose-6-phosphat
7	32	94.1	569	2	S23542	glucose-6-phosphat
8	32	94.1	569	2	S41806	glucose-6-phosphat
9	32	94.1	569	2	S57831	glucose-6-phosphat
10	32	94.1	1295	2	S60179	pol polyprotein ho
11	31	91.2	85	2	F75291	hypothetical prote
12	31	91.2	105	2	T38761	hypothetical prote
13	31	91.2	158	2	AI0605	conserved hypothet
14	31	91.2	158	2	E64823	hypothetical prote
15	31	91.2	158	2	E90745	probable sensory t
16	31	91.2	287	2	C83872	hypothetical prote
17	31	91.2	299	2	B82535	conserved hypothet
18	31	91.2	384	2	S34816	nitrogenase cofact
19	31	91.2	409	2	S01825	transforming growt
20	31	91.2	410	2	A41397	transforming growt
21	31	91.2	410	2	A55706	transforming growt
22	31	91.2	412	2	A34939	transforming growt
23	31	91.2	412	2	A36169	transforming growt
24	31	91.2	415	1	O4PSCP	camphor 5-monooxyg
25	31	91.2	505	2	T19971	hypothetical prote
26	31	91.2	506	2	T19973	hypothetical prote
27	31	91.2	554	2	A96018	probable regulator
28	31	91.2	639	2	A56126	peroxisomal target
29	31	91.2	936	2	E96508	hypothetical prote

30	91.2	984	2	T48216	hypothetical prote
31	91.2	1094	2	F70697	probable arabinosy
32	91.2	1630	2	T00390	KIAA0614 protein -
33	91.2	4342	2	H83343	probable non-ribos
34	88.2	120	2	S10865	early E4 13K prote
35	88.2	158	2	D72305	hypothetical prote
36	88.2	181	2	G89799	hypothetical prote
37	88.2	233	2	S52461	hypothetical prote
38	88.2	233	2	S38228	hypothetical prote
39	88.2	233	2	S53104	hypothetical prote
40	88.2	239	2	E69189	hypothetical prote
41	88.2	263	2	A30227	hypothetical prote
42	88.2	284	2	F83042	hypothetical prote
43	88.2	292	2	F86431	hypothetical prote
44	88.2	308	2	S77938	EBNA-LP protein -
45	88.2	319	2	F82601	5'-nucleotidase XF
46	88.2	332	2	A29711	deacetoxycephalosp
47	88.2	338	2	E90998	hypothetical prote
48	88.2	346	2	D85818	unknown protein en
49	88.2	394	2	AF2963	phosphomannose iso
50	88.2	394	2	G98319	mannose-6-phosphat
51	88.2	406	2	A39339	protein C inhibito
52	88.2	427	2	B83001	probable C4-dicarb
53	88.2	497	2	T15812	hypothetical prote
54	88.2	515	2	G70941	hypothetical prote
55	88.2	521	2	T11166	CDPdiacylglycerol-
56	88.2	521	2	T01923	hypothetical prote
57	88.2	658	1	A65079	hypothetical prote
58	88.2	658	2	AG0876	arginine decarboxy
59	88.2	658	2	F91105	arginine decarboxy
60	88.2	658	2	A85951	biosynthetic argin
61	88.2	659	2	AB0114	biosynthetic argin
62	88.2	952	2	E86147	arginine decarboxy
63	88.2	1040	2	T08190	T1N6.4 protein - A
64	88.2	1094	2	C70612	hypothetical prote
65	88.2	1123	2	T28139	probable recB prot
66	88.2	6260	2	T30228	PK4 protein kinase
67	88.2	8563	2	T30226	polyketide synthas
68	85.3	65	2	S59524	polyketide synthas
69	85.3	80	2	H82194	triose-phosphate i
70	85.3	84	2	T03339	conserved hypothet
71	85.3	106	1	Z2BPT9	gene e28 protein -
72	85.3	114	2	AI0294	gene 49.2 protein
73	85.3	124	2	AG3149	probable membrane
74	85.3	124	2	E98138	hypothetical prote
75	85.3	143	2	E71041	hypothetical prote
76	85.3	165	2	A70326	conserved hypothet
77	85.3	192	2	B81911	probable molybdopt
78	85.3	199	2	T03101	hypothetical prote
79	85.3	201	2	G59096	hypothetical prote
80	85.3	232	2	B82987	probable hydrolase
81	85.3	243	2	S38236	hypothetical prote
82	85.3	251	2	T35258	probable transcrip
83	85.3	255	2	F64503	hypothetical prote
84	85.3	269	2	S73999	hypothetical prote
85	85.3	277	2	G71456	hypothetical prote
86	85.3	283	2	A64174	hypothetical prote
87	85.3	283	2	T31147	hypothetical prote
88	85.3	285	2	H72073	phospho-2-dehydro-
89	85.3	285	2	H86550	deoxyheptonate ald
90	85.3	289	2	E83454	cell division prot
91	85.3	301	2	C82394	probable glycine c
92	85.3	308	2	S58504	reverse transcript
93	85.3	327	2	D82090	conserved hypothet
94	85.3	337	2	F91191	probable LPS biosy
95	85.3	337	2	G86038	probable LPS biosy
96	85.3	338	2	AD1142	tagatose-1,6-dipho
97	85.3	338	2	AG1500	tagatose-1,6-dipho
98	85.3	339	2	C86978	hypothetical prote
99	85.3	339	2	F70901	riboflavin bifunct
100	85.3	341	2	E65080	hypothetical prote
101	85.3	341	2	B91107	probable protein t
102	85.3	341	2	E85952	probable protein t

103 29 85.3 344 2 E84377 protein export [im
104 29 85.3 350 2 AG0879 A/G-specific adeni
105 29 85.3 350 2 A40647 methylation-indepe
106 29 85.3 353 2 AI0036 conserved hypotet
107 29 85.3 357 2 E83205 hypotetrical prote
108 29 85.3 359 1 F64407 carotenoid biosynt
109 29 85.3 364 1 C69351 probable iron-sulf
110 29 85.3 369 2 C75043 hypotetrical prote
111 29 85.3 370 2 AB2578 membrane-bound lyt
112 29 85.3 370 2 A97360 outer membrane lip
113 29 85.3 394 2 C82439 peptide methionine
114 29 85.3 399 2 D96832 hypotetrical prote
115 29 85.3 414 2 AI2823 hypotetrical prote
116 29 85.3 416 2 E69374 aminotransferase l
117 29 85.3 416 2 F81337 RNA polymerase sig
118 29 85.3 417 1 S67566 probable membrane
119 29 85.3 423 2 E86132 hypotetrical prote
120 29 85.3 423 2 C91291 hypotetrical prote
121 29 85.3 426 2 S56561 hypotetrical 48.2K
122 29 85.3 440 2 A83435 ATP synthase in ty
123 29 85.3 441 2 C64026 hypotetrical prote
124 29 85.3 454 2 I64124 virion morphogenes
125 29 85.3 482 2 T50901 light-harvesting l
126 29 85.3 485 2 E83620 probable amidase p
127 29 85.3 540 2 B45665 adult-specific 61.
128 29 85.3 541 2 AF2483 hypotetrical prote
129 29 85.3 567 2 T03948 probable glucose-6
130 29 85.3 567 2 T02094 hypotetrical prote
131 29 85.3 570 2 S09812 hypotetrical prote
132 29 85.3 576 2 H70961 CTP synthase (EC 6
133 29 85.3 597 2 T00471 dopamine transport
134 29 85.3 619 2 S20346 dopamine transport
135 29 85.3 619 2 I59558 dopamine transport
136 29 85.3 620 2 I57937 dopamine transport
137 29 85.3 633 2 B70946 dopamine transport
138 29 85.3 633 2 C70742 NADH2 dehydrogenas
139 29 85.3 713 2 JW0088 probable rsbU prot
140 29 85.3 750 2 A97501 3',5'-cyclic-nucle
141 29 85.3 750 2 AE2719 topoisomerase iv c
142 29 85.3 775 2 AB3402 topoisomerase iv c
143 29 85.3 816 2 A71006 hypotetrical prote
144 29 85.3 860 2 C82750 mannosyltransferas
145 29 85.3 897 2 T21688 hypotetrical prote
146 29 85.3 924 2 E71476 alanine-tRNA ligas
147 29 85.3 1003 2 T34066 hypotetrical prote
148 29 85.3 1019 2 JC7538 neuronal different
149 29 85.3 1092 2 H69071 DNA-directed DNA p
150 29 85.3 1150 2 S49956 probable membrane
151 29 85.3 1204 2 T18812 hypotetrical prote
152 29 85.3 1252 2 D71810 probable type II D
153 29 85.3 1279 2 E64709 type IIS restricti
154 29 85.3 1369 2 T03104 tegument protein h
155 29 85.3 1545 2 T26589 hypotetrical prote
156 29 85.3 1774 2 S13178 6-methylsalicylic
157 29 85.3 2225 1 A23443 pyrimidine synthes
158 29 85.3 55 2 T30476 hypotetrical prote
159 29 85.3 105 2 PW0018 hypotetrical prote
160 29 85.3 105 2 D72546 hypotetrical prote
161 29 85.3 120 1 QQXRS3 nonstructural prot
162 29 85.3 140 2 AD1205 molybdopterin conv
163 29 85.3 140 2 AC1562 molybdopterin conv
164 29 85.3 147 2 H82172 heat shock protein
165 29 85.3 182 2 AH0227 CDPdiacylglycerol-
166 29 85.3 209 1 XURTMC methylated-DNA-[pr
167 29 85.3 244 2 A82316 hypotetrical prote
168 29 85.3 251 2 AB3390 arginine/ornithine
169 29 85.3 258 2 T08265 conserved hypotet
170 29 85.3 267 2 T36675 probable integral
171 29 85.3 279 2 H81678 phospho-2-dehydro-
172 29 85.3 317 2 A72073 hypotetrical prote
173 29 85.3 317 2 F86552 hypotetrical prote

176 28 82.4 339 2 AG0799 probable transcrip
177 28 82.4 339 2 S77001 hypotetrical prote
178 28 82.4 344 2 H70030 conserved hypotet
179 28 82.4 373 2 PW0042 hypotetrical prote
180 28 82.4 401 2 F87196 activin - fruit fl
181 28 82.4 408 2 T50876 probable membrane-
182 28 82.4 413 2 A84337 hypotetrical membr
183 28 82.4 426 2 A83330 hypotetrical prote
184 28 82.4 436 2 F86486 probable RND efflu
185 28 82.4 441 2 H82617 protein F28J9.3 [i
186 28 82.4 452 2 F83587 folylpolyglutamate
187 28 82.4 524 2 E71881 inner membrane pro
188 28 82.4 527 2 B64633 hypotetrical prote
189 28 82.4 557 2 T44843 glucose-6-phosphat
190 28 82.4 600 2 T24626 hypotetrical prote
191 28 82.4 610 2 A57632 homeotic protein B
192 28 82.4 636 2 AI2552 transposase alr801
193 28 82.4 663 2 B84194 DNA topoisomerase
194 28 82.4 678 2 AB2744 two component sens
195 28 82.4 691 2 T44543 probable bacteriop
196 28 82.4 713 2 G98286 methylmalonyl-CoA
197 28 82.4 725 2 AI1544 conserved hypotet
198 28 82.4 726 2 G83310 conserved hypotet
199 28 82.4 741 2 D83633 conserved hypotet
200 28 82.4 748 2 S59327 hypotetrical prote
201 28 82.4 767 2 AE3370 sensory transducti
202 28 82.4 821 2 B71229 pyruvate, water di
203 28 82.4 856 2 E70875 hypotetrical prote
204 28 82.4 861 2 A97525 cell cycle histidi
205 28 82.4 863 2 E87095 conserved hypotet
206 28 82.4 886 1 JC5085 replication licens
207 28 82.4 904 2 C70559 probable polA prot
208 28 82.4 914 2 E83901 hypotetrical prote
209 28 82.4 957 2 A47531 glutamyl aminopept
210 28 82.4 1494 2 T13798 hypotetrical prote
211 28 82.4 1570 2 AC2012 hypotetrical prote
212 28 82.4 1583 2 F86366 protein F26F24.8 l
213 28 82.4 2391 2 T18410 carbamoyl-phosphat
214 28 82.4 73 2 H81298 probable molybdopt
215 28 82.4 78 2 B90830 excisionase [impor
216 27 79.4 78 2 H85687 probable integrase
217 27 79.4 95 2 AH2789 hypotetrical prote
218 27 79.4 95 2 G97568 hypotetrical prote
219 27 79.4 96 2 S54425 orotate phosphorib
220 27 79.4 98 2 A64143 hypotetrical prote
221 27 79.4 101 2 G97567 hypotetrical prote
222 27 79.4 101 2 AF2788 hypotetrical prote
223 27 79.4 102 2 C96034 conserved hypotet
224 27 79.4 105 2 AI0557 BOLA protein [limpo
225 27 79.4 114 1 Q4ADD2 early E4 13K prote
226 27 79.4 117 2 T37543 60s ribosomal prot
227 27 79.4 121 2 C84353 hypotetrical prote
228 27 79.4 124 2 D83105 transcription regu
229 27 79.4 125 2 T43208 ribosomal protein
230 27 79.4 130 2 T15648 hypotetrical prote
231 27 79.4 133 2 JC2202 secretin precursor
232 27 79.4 136 2 G72573 hypotetrical prote
233 27 79.4 137 2 S64094 hypotetrical prote
234 27 79.4 139 2 G71978 hypotetrical prote
235 27 79.4 139 2 H64527 hypotetrical prote
236 27 79.4 140 2 T26470 alpha-lactalbumin
237 27 79.4 142 1 LAGP hypotetrical prote
238 27 79.4 151 2 T38090 hypotetrical prote
239 27 79.4 156 2 A69260 molybdopterin bios
240 27 79.4 157 2 B69659 probable sensory t
241 27 79.4 161 2 T06826 beta-fructofuranos
242 27 79.4 162 2 A34429 dihydrofolate redu
243 27 79.4 174 2 AB1888 hypotetrical prote
244 27 79.4 178 2 S77189 hypotetrical prote
245 27 79.4 182 1 XNECPG CDPdiacylglycerol-

249 27 79.4 182 2 AC0749 CDPdiacylglycerol-
250 27 79.4 182 2 B90960 CDPdiacylglycerol-
251 27 79.4 182 2 C85808 CDPdiacylglycerol-
252 27 79.4 182 2 AG2062 hypotheical prote
253 27 79.4 185 2 H82227 CDP-diacylglycerol
254 27 79.4 185 2 F64049 CDPdiacylglycerol-
255 27 79.4 186 2 B83322 CDP-diacylglycerol
256 27 79.4 187 2 C81803 CDPdiacylglycerol-
257 27 79.4 187 2 F81065 CDP-diacylglycerol
258 27 79.4 187 2 A82746 conserved hypothet
259 27 79.4 189 2 A12653 molybdenum cofacto
260 27 79.4 189 2 E97435 molybdopterin bios
261 27 79.4 189 2 A83586 probable transcrip
262 27 79.4 196 2 S40728 hypotheical prote
263 27 79.4 198 2 S73275 hypotheical prote
264 27 79.4 206 1 I40173 orotate phosphorib
265 27 79.4 206 2 AD1898 urease accessory p
266 27 79.4 207 2 G97633 thymidine kinase (
267 27 79.4 209 2 AG1303 orotate phosphorib
268 27 79.4 209 2 AG1675 orotate phosphorib
269 27 79.4 211 2 G82573 CDP-diacylglycerol
270 27 79.4 213 1 E64798 Nicotinate-nucleot
271 27 79.4 213 2 A85564 hypotheical prote
272 27 79.4 213 2 E90713 hypotheical prote
273 27 79.4 215 2 T18842 hypotheical prote
274 27 79.4 216 2 H69953 conserved hypothet
275 27 79.4 216 2 T18176 hypotheical prote
276 27 79.4 221 2 A82190 hypotheical prote
277 27 79.4 226 2 H82710 phosphoglycolate p
278 27 79.4 227 2 T49725 hypotheical prote
279 27 79.4 230 2 H86758 hypotheical prote
280 27 79.4 231 2 E81215 thiol-disulfide in
281 27 79.4 234 2 T46203 transcription fact
282 27 79.4 241 2 F95355 protein limported
283 27 79.4 243 2 B72757 probable phosphate
284 27 79.4 244 2 T07068 beta-fructofuranos
285 27 79.4 246 2 B55545 int protein - Esch
286 27 79.4 246 2 T00231 recombinase - Esch
287 27 79.4 247 2 S18604 triose-phosphate i
288 27 79.4 248 2 AE1040 conserved hypothet
289 27 79.4 248 2 S56531 hypotheical 27.4K
290 27 79.4 248 2 A12175 hypotheical prote
291 27 79.4 253 1 ISZMT triose-phosphate i
292 27 79.4 260 2 G83989 hydrolase BH2719 [
293 27 79.4 261 2 T32399 hypotheical prote
294 27 79.4 262 2 C75554 conserved hypothet
295 27 79.4 267 2 S55487 probable 3-methyl-
296 27 79.4 270 2 A98141 hypotheical prote
297 27 79.4 270 2 A13146 glucose-1-phosphat
298 27 79.4 275 2 B81430 probable prephenat
299 27 79.4 283 2 A86523 probable glycosyl
300 27 79.4 288 2 S73016 H+-transporting tw
301 27 79.4 290 2 B97100 pyridoxal kinase r
302 27 79.4 292 2 F84951 hypotheical prote
303 27 79.4 293 2 T00466 coatomer complex e
304 27 79.4 294 2 T10858 probable carboxype
305 27 79.4 295 2 AB3613 glucose-1-phosphat
306 27 79.4 296 2 AG0147 probable membrane
307 27 79.4 297 1 JN0773 calponin H1 - rat
308 27 79.4 297 1 S31484 calponin H1 - mous
309 27 79.4 297 1 S31486 calponin H1 - mous
310 27 79.4 297 2 JC4500 basic calponin - h
311 27 79.4 297 2 G02142 smooth muscle cell
312 27 79.4 297 2 S55063 conserved hypothet
313 27 79.4 299 2 A75401 ribulose-phosphate
314 27 79.4 301 2 S62416 n-acetylglucosamin
315 27 79.4 302 2 T36030 probable fructokin
316 27 79.4 309 2 S23096 protein kinase (EC
317 27 79.4 310 2 B55171 chloroplast import
318 27 79.4 315 2 T18696 hypotheical prote
319 27 79.4 315 2 B98226 hypotheical prote
320 27 79.4 315 2 AF3060 conserved hypothet
321 27 79.4 320 2 AG0934 bifunctional prote

322 27 79.4 321 1 BVECBF
323 27 79.4 321 2 A86089 hypotheical prote
324 27 79.4 321 2 D91241 hypotheical prote
325 27 79.4 324 2 C82251 GGDEF family prote
326 27 79.4 325 1 G90393 phosphoesterase-re
327 27 79.4 325 2 T04962 catechol O-methylt
328 27 79.4 329 2 T19872 hypotheical prote
329 27 79.4 330 2 F82338 hypotheical prote
330 27 79.4 335 2 D71275 hypotheical prote
331 27 79.4 336 2 A64378 hypotheical prote
332 27 79.4 336 2 T26189 hypotheical prote
333 27 79.4 337 2 E69136 hypotheical prote
334 27 79.4 337 2 T28690 hypotheical prote
335 27 79.4 345 2 D88485 hypotheical prote
336 27 79.4 350 2 B38535 protein F23F12.4 [
337 27 79.4 350 2 H85953 A/G-specific adeni
338 27 79.4 350 2 E91108 adenine glycosylas
339 27 79.4 360 2 S54256 photosystem II pro
340 27 79.4 361 1 F65012 hypotheical prote
341 27 79.4 361 2 A91037 probable peptidase
342 27 79.4 361 2 C85881 probable peptidase
343 27 79.4 361 2 T25196 hypotheical prote
344 27 79.4 363 2 E90268 acyl-CoA dehydroge
345 27 79.4 365 2 T25309 hypotheical prote
346 27 79.4 367 2 T43252 carbamoyl-phosphat
347 27 79.4 369 2 T40905 WD repeat protein
348 27 79.4 376 2 T51170 homocitrate syntha
349 27 79.4 376 2 T35868 probable dipeptida
350 27 79.4 376 2 A75449 phospho-2-dehydro-
351 27 79.4 377 2 C70628 hypotheical prote
352 27 79.4 378 2 F90323 coenzyme PQQ synth
353 27 79.4 378 2 D90347 coenzyme PQQ synth
354 27 79.4 379 2 JQ2272 formate dehydrogen
355 27 79.4 379 2 F71533 hypotheical prote
356 27 79.4 382 2 T04963 catechol O-methylt
357 27 79.4 383 2 S32457 gamma-butyrobetain
358 27 79.4 384 2 E81689 conserved hypothet
359 27 79.4 385 1 T52057 trehalose-phosphat
360 27 79.4 389 2 B69096 corrinoid/iron-sul
361 27 79.4 391 2 A71936 hypotheical prote
362 27 79.4 391 2 B64648 conserved hypothet
363 27 79.4 392 2 E72219 conserved hypothet
364 27 79.4 396 2 T45088 pyruvate synthase
365 27 79.4 397 2 D69174 3-hydroxy-3-methyl
366 27 79.4 404 2 B64543 hypotheical prote
367 27 79.4 409 2 D75629 probable transposa
368 27 79.4 413 2 JV0087 glucose-1-phosphat
369 27 79.4 413 2 F90773 periplasmic glucos
370 27 79.4 413 2 B85636 periplasmic glucos
371 27 79.4 414 2 C75461 hypotheical prote
372 27 79.4 416 2 D75407 probable transposa
373 27 79.4 417 2 S07286 hypotheical prote
374 27 79.4 420 2 B87201 hypotheical prote
375 27 79.4 420 2 A70796 probable cycloprop
376 27 79.4 420 2 E75378 probable valine-py
377 27 79.4 423 2 A55797 collagen precursor
378 27 79.4 424 2 S63583 tetrahydrofolylpol
379 27 79.4 424 2 B81936 tetrahydrofolylpol
380 27 79.4 424 2 C81170 folylpolyglutamate
381 27 79.4 431 2 JW0098 carbazole dioxigen
382 27 79.4 433 2 C75354 conserved hypothet
383 27 79.4 436 2 D69270 ferredoxin-nitrite
384 27 79.4 436 2 A38145 invariant surface
385 27 79.4 454 2 AG1221 glutamate-cysteine
386 27 79.4 454 2 AB1575 ethanolamine ammon
387 27 79.4 454 2 A97048 ethanolamine ammon
388 27 79.4 464 2 H83006 similar to phospho
389 27 79.4 466 2 JC6525 hypotheical prote
390 27 79.4 468 2 H70521 processing peptida
391 27 79.4 468 2 AD1926 probable papa2 pro
392 27 79.4 469 2 AE2188 hypotheical prote
393 27 79.4 469 2 D82575 hypotheical prote
394 27 79.4 470 2 D82575 phage-related prot

395	27	79.4	471	2	B38637	Ras inhibitor (clo
396	27	79.4	472	2	E82545	phage-related prot
397	27	79.4	473	2	T06167	beta-fructofuranos
398	27	79.4	481	2	T35149	two-component syst
399	27	79.4	485	2	S56359	probable proton/ol
400	27	79.4	485	2	F86108	probable peptide t
401	27	79.4	485	2	H91267	probable peptide t
402	27	79.4	499	2	AC2068	cell death suppress
403	27	79.4	505	2	E83918	carboxypeptidase B
404	27	79.4	509	2	T30861	traJ protein - Sal
405	27	79.4	513	1	S50216	translation initia
406	27	79.4	540	2	S76869	hypothetical prote
407	27	79.4	542	2	AB1066	conserved hypothet
408	27	79.4	553	2	D83640	hypothetical prote
409	27	79.4	555	2	F71083	hypothetical prote
410	27	79.4	555	2	A43784	Xpo protein - Afri
411	27	79.4	562	2	F72771	probable lysyl-trn
412	27	79.4	575	2	C86398	protein T7N9.26 [i
413	27	79.4	581	2	T12095	beta-fructofuranos
414	27	79.4	581	2	F84599	probable kinesin h
415	27	79.4	584	2	S51882	topoisomerase I-re
416	27	79.4	590	2	D69722	thiamin biosynthes
417	27	79.4	592	2	S56681	beta-fructofuranos
418	27	79.4	592	2	S25705	Ig mu chain - shee
419	27	79.4	595	2	E83891	thiamin biosynthes
420	27	79.4	595	2	T39228	beta-transducin -
421	27	79.4	606	1	UZAD12	terminal protein p
422	27	79.4	611	2	F75095	probable asparagin
423	27	79.4	618	2	B87564	conserved hypothet
424	27	79.4	623	2	T48859	disease resistance
425	27	79.4	625	2	S60401	nuclear division p
426	27	79.4	634	1	GERTX1	matrix glycoprotei
427	27	79.4	655	2	T26061	hypothetical prote
428	27	79.4	664	2	S60062	hevin precursor -
429	27	79.4	670	2	S76872	hypothetical prote
430	27	79.4	676	1	A39379	hatching-suppresse
431	27	79.4	686	2	F89895	ATP-dependent DNA
432	27	79.4	693	2	AE0005	ATP-dependent DNA
433	27	79.4	716	1	BVECAD	membrane protein t
434	27	79.4	738	2	S10659	membrane protein t
435	27	79.4	757	2	B75437	ABC transporter, A
436	27	79.4	763	2	E96571	hypothetical prote
437	27	79.4	788	1	B4AG58	virB4 protein prec
438	27	79.4	789	1	B4AGA6	virB4 protein prec
439	27	79.4	789	2	AH3248	component of type
440	27	79.4	795	2	H84944	phenylalanine-tRNA
441	27	79.4	798	2	F69549	acetyl-CoA decarbo
442	27	79.4	803	2	E82392	phosphoenolpyruvat
443	27	79.4	805	1	A55054	calpain (EC 3.4.22
444	27	79.4	805	2	A39764	carbon-monoxide de
445	27	79.4	809	2	B83409	hypothetical prote
446	27	79.4	822	2	AD3232	conjugal transfer
447	27	79.4	834	2	T06055	hypothetical prote
448	27	79.4	848	2	C81356	aconitate hydratase
449	27	79.4	850	2	C83081	probable oxidoredu
450	27	79.4	852	2	F44020	hypothetical prote
451	27	79.4	852	2	T08502	trbE protein - Ent
452	27	79.4	906	2	G96621	probable disease r
453	27	79.4	968	2	T45746	hypothetical prote
454	27	79.4	978	2	T14968	phage lambda-relat
455	27	79.4	1009	2	S44621	C50C3.2 protein -
456	27	79.4	1014	2	JE0333	klotho protein - r
457	27	79.4	1047	2	S19508	MSH3 protein - yea
458	27	79.4	1056	2	T28636	zinc-metalloprotein
459	27	79.4	1173	2	H83189	DNA polymerase III
460	27	79.4	1175	2	T46124	hypothetical prote
461	27	79.4	1221	2	T18550	reverse gyrase cha
462	27	79.4	1272	2	C90593	hypothetical prote
463	27	79.4	1349	2	T13031	nucleoporin - frui
464	27	79.4	1365	2	T13991	nucleoporin 154 -
465	27	79.4	1541	1	S71839	canalicular multidi
466	27	79.4	1608	1	WMTMG	183K protein - tob
467	27	79.4	1899	2	T32732	PAM C-terminal int

27	79.4	1919	2	T42098	PAM interacting pr
27	79.4	2137	2	T05244	hypothetical prote
27	79.4	2157	2	AI3009	peptide synthetase
27	79.4	2186	2	T13169	tiggrin - fruit fl
27	79.4	2233	1	ZLN2P3	genome polyprotein
27	79.4	2255	1	JQ1532	genome polyprotein
27	79.4	2298	2	T49648	hypothetical prote
27	79.4	2326	2	T29140	hypothetical prote
27	79.4	2566	2	E98274	hypothetical prote
27	79.4	2761	2	T29285	hypothetical prote
27	79.4	3163	2	AB0233	yersiniabactin bio
27	79.4	3163	2	T17440	probable polyketid
27	79.4	4196	2	T43274	dynein heavy chain
27	79.4	15281	2	S41309	cyclosporin synthe
26	76.5	35	2	H64642	hypothetical prote
26	76.5	70	2	T06920	ribosomal protein
26	76.5	72	1	RSBPXL	excisionase - phag
26	76.5	72	2	A90729	excisionase [impor
26	76.5	72	2	S06533	excisionase - phag
26	76.5	72	2	C70241	hypothetical prote
26	76.5	73	2	H71893	probable molybdopt
26	76.5	74	2	A64620	molybdopterin conv
26	76.5	82	1	CCPS5D	cytochrome c551 -
26	76.5	87	2	S42171	cytochrome-c oxida
26	76.5	87	2	T10462	hypothetical prote
26	76.5	96	2	H87336	conserved hypothet
26	76.5	101	2	F72573	hypothetical prote
26	76.5	103	2	T07053	cysteine proteinas
26	76.5	106	2	JQ0234	hypothetical 12.5K
26	76.5	107	2	S50516	hypothetical prote
26	76.5	108	1	RPECW	trp operon repress
26	76.5	108	1	S21331	hypothetical prote
26	76.5	108	2	S45254	trp operon repress
26	76.5	108	2	G91297	regulator for trp
26	76.5	108	2	A86139	trp operon repress
26	76.5	109	2	S45257	trp operon repress
26	76.5	121	2	A70555	hypothetical prote
26	76.5	123	1	LACM	alpha-lactalbumin
26	76.5	125	2	JU0179	heparin-binding pr
26	76.5	125	2	AC0610	heparin-binding pr
26	76.5	127	2	E86158	probable integrase
26	76.5	128	1	S52084	60S ribosomal prot
26	76.5	128	2	JC2120	ribosomal protein
26	76.5	128	2	JC2121	heparin-binding pr
26	76.5	128	2	JC2119	heparin-binding pr
26	76.5	130	2	A30033	development-specif
26	76.5	130	2	I83571	probable membrane
26	76.5	130	2	C85701	hypothetical prote
26	76.5	130	2	F90843	hypothetical prote
26	76.5	131	2	S74539	hypothetical prote
26	76.5	133	2	S45752	hypothetical prote
26	76.5	137	2	A87586	hypothetical prote
26	76.5	137	2	B96603	transcription fact
26	76.5	138	2	C84068	polyribonucleotide
26	76.5	138	2	T40987	hypothetical prote
26	76.5	142	2	A82966	conserved hypothet
26	76.5	143	2	A70953	hypothetical prote
26	76.5	143	2	C75173	hypothetical prote
26	76.5	144	2	F97571	hypothetical prote
26	76.5	144	2	AG2792	conserved hypothet
26	76.5	148	2	S26450	hypothetical prote
26	76.5	153	2	AD3339	transporter BMEI06
26	76.5	154	2	E75024	hypothetical prote
26	76.5	156	2	C48230	hypothetical prote
26	76.5	156	2	H84251	T-cell receptor be
26	76.5	157	2	B69476	flagella-related p
26	76.5	158	2	S43892	inosine monophosph
26	76.5	158	2	G64749	DNA-directed DNA p
26	76.5	161	2	A69732	hypothetical prote
26	76.5	162	2	G70005	PBSX prophage ORF
26	76.5	163	2	T36985	hypothetical prote
26	76.5	164	2	AF3220	conserved hypothet
26	76.5	166	2	H64416	VPS29-like phospho

541 26 76.5 166 2 T40988 probable glutaredo
542 26 76.5 174 2 G83712 hypothetical prote
543 26 76.5 175 2 D85544 primosomal replica
544 26 76.5 175 2 H90693 primosomal replica
545 26 76.5 175 2 JQ1149 conserved hypotHet
546 26 76.5 176 2 B64549 hypothetical prote
547 26 76.5 176 2 F83531 hypothetical prote
548 26 76.5 177 2 AB2294 hypothetical prote
549 26 76.5 177 2 H87412 hypothetical prote
550 26 76.5 178 2 E91080 hypothetical prote
551 26 76.5 178 2 F85925 hypothetical prote
552 26 76.5 180 2 G70028 conserved hypotHet
553 26 76.5 183 2 AG1003 conserved oxidoredu
554 26 76.5 183 2 AI0023 conserved hypotHet
555 26 76.5 184 1 B65129 probable NAD(P)H2
556 26 76.5 184 2 G85999 probable NAD(P)H o
557 26 76.5 184 2 B91154 probable NAD(P)H o
558 26 76.5 184 2 H72414 hypothetical prote
559 26 76.5 184 2 H84143 hypothetical prote
560 26 76.5 185 2 E70029 conserved hypotHet
561 26 76.5 185 2 D82838 conserved hypotHet
562 26 76.5 186 2 G91158 hypothetical prote
563 26 76.5 186 2 E86004 hypothetical prote
564 26 76.5 186 2 H65134 hypothetical 21.2
565 26 76.5 187 2 C82065 hypothetical prote
566 26 76.5 188 2 AD0533 conserved hypotHet
567 26 76.5 189 2 B69020 hypothetical prote
568 26 76.5 190 2 A82343 conserved hypotHet
569 26 76.5 191 1 H64744 yaed protein - Esc
570 26 76.5 191 2 F84932 hypothetical prote
571 26 76.5 191 2 B85505 probable phosphata
572 26 76.5 191 2 B90654 probable phosphata
573 26 76.5 192 2 G95072 conserved hypotHet
574 26 76.5 192 2 D97940 conserved hypotHet
575 26 76.5 193 2 A86852 hypothetical prote
576 26 76.5 198 2 B84382 hypothetical prote
577 26 76.5 198 2 AD0999 probable NUPix hyd
578 26 76.5 199 2 E85635 part of trimethyla
579 26 76.5 199 2 A90773 part of trimethyla
580 26 76.5 200 2 T23485 hypothetical prote
581 26 76.5 202 2 AC2875 hypothetical prote
582 26 76.5 203 2 S57836 lymphoid-specific
583 26 76.5 204 2 E90668 probable membrane
584 26 76.5 204 2 H85518 probable membrane
585 26 76.5 204 2 G64754 hypothetical prote
586 26 76.5 205 2 T16560 probable glutathio
587 26 76.5 207 2 T37464 conserved hypotHet
588 26 76.5 208 2 AD2808 hypothetical prote
589 26 76.5 209 2 B87627 hydrolase, Cbby/Cb
590 26 76.5 211 2 B75470 uracil-DNA glycosy
591 26 76.5 213 2 AC3389 hypothetical prote
592 26 76.5 215 2 E96533 hypothetical prote
593 26 76.5 217 2 B97587 hypothetical prote
594 26 76.5 217 2 T35633 hypothetical prote
595 26 76.5 218 2 B71183 conserved hypotHet
596 26 76.5 218 2 D72073 CT383 hypothetical
597 26 76.5 218 2 D86550 hypothetical prote
598 26 76.5 218 2 A82765 phosphoglycolate p
599 26 76.5 219 2 E75143 protein T23J18.24
600 26 76.5 223 2 A86249 conserved hypotHet
601 26 76.5 225 2 D83200 ABC transporter, A
602 26 76.5 228 2 C87467 carboxylesterase (
603 26 76.5 228 2 AI3547 hypothetical prote
604 26 76.5 229 2 S64049 conserved hypotHet
605 26 76.5 236 2 D72355 glutathione transf
606 26 76.5 236 2 AF3416 5'TG3'-interactin
607 26 76.5 237 2 JC7392 infected cell prot
608 26 76.5 246 2 B72771 hypothetical prote
609 26 76.5 248 1 JQ1682 beta-fructofuranos
610 26 76.5 249 2 H86786 triose-phosphate i
611 26 76.5 250 2 T14548 esterase - Thermot
612 26 76.5 253 2 S53760
613 26 76.5 253 2 E72302

614 26 76.5 253 2 C84644
615 26 76.5 254 1 TRW3Y
616 26 76.5 257 2 C70715
617 26 76.5 260 2 S17300
618 26 76.5 260 2 S14157
619 26 76.5 260 2 A82837 conserved hypotHet
620 26 76.5 260 2 T41554 hypothetical prote
621 26 76.5 262 2 AI2480 hypothetical prote
622 26 76.5 262 2 C91111 probable tropinone
623 26 76.5 263 2 F84694 endo-1,3-1,4-beta-
624 26 76.5 263 2 AB3055 trypsin-like prote
625 26 76.5 264 2 S32794 probable tropinone
626 26 76.5 267 2 AC3133 hypothetical prote
627 26 76.5 268 2 E84695 hypothetical prote
628 26 76.5 269 2 C84314 hypothetical prote
629 26 76.5 269 2 H84378 transcription fact
630 26 76.5 269 2 S58439 probable transcrip
631 26 76.5 274 2 E83444 hypothetical prote
632 26 76.5 274 2 T26658 hypothetical prote
633 26 76.5 276 2 B70598 PBX prophage prot
634 26 76.5 277 2 I40410 DNA-directed DNA p
635 26 76.5 278 2 S06416 probable deoxyhept
636 26 76.5 278 2 B71522 hypothetical prote
637 26 76.5 279 2 T26166 hypothetical prote
638 26 76.5 279 2 T21068 hypothetical prote
639 26 76.5 279 2 T15662 lignin beta-ether
640 26 76.5 281 2 A43749 probable transposa
641 26 76.5 281 2 F85956 hypothetical prote
642 26 76.5 286 2 G70537 endo-1,3-1,4-beta-
643 26 76.5 289 2 C98231 isochorismatase (E
644 26 76.5 290 2 AB3519 phosphoribosylamin
645 26 76.5 291 1 JH0489 phosphoribosylamin
646 26 76.5 291 2 S55291 phosphoribosylamin
647 26 76.5 291 2 S55292 phosphoribosylamin
648 26 76.5 293 2 S22613 abequose synthase
650 26 76.5 294 2 AC1420 xylose operon regu
651 26 76.5 294 2 AD1795 xylose operon regu
652 26 76.5 294 2 A96155 hypothetical prote
653 26 76.5 295 2 H91117 probable transcrip
654 26 76.5 295 2 G85962 probable transcrip
655 26 76.5 301 2 JC2039 phosphoribosylamin
656 26 76.5 306 1 JQ1395 conserved hypotHet
657 26 76.5 307 2 AB0562 conserved hypotHet
658 26 76.5 308 2 AC1867 hypothetical prote
659 26 76.5 311 2 G64238 hypothetical prote
660 26 76.5 311 2 AG3526 2-hydroxyhepta-2,4
661 26 76.5 316 2 D83406 probable transmemb
662 26 76.5 316 2 T32962 hypothetical prote
663 26 76.5 317 2 E95239 conserved hypotHet
664 26 76.5 317 2 E98103 conserved hypotHet
665 26 76.5 317 2 G83544 probable transcrip
666 26 76.5 321 2 S58686 probable transcrip
667 26 76.5 322 2 H87131 probable lytic mur
668 26 76.5 323 2 G82135 pectic enzyme secr
669 26 76.5 324 2 D37874 ribosomal large ch
670 26 76.5 324 2 G82289 hypothetical prote
671 26 76.5 326 2 S73828 probable transposa
672 26 76.5 327 2 E82277 dTDPglucose 4,6-de
673 26 76.5 328 1 DWSMGG replication protei
674 26 76.5 328 2 A35257 hypothetical prote
675 26 76.5 331 2 B90173 hypothetical prote
676 26 76.5 331 2 S45584 transcription acti
677 26 76.5 332 2 B47017 probable transcrip
678 26 76.5 332 2 AD2541 transcription init
679 26 76.5 333 2 E75332 probable ADP-ribos
680 26 76.5 333 2 E97257 spore coat protein
681 26 76.5 334 2 I59348 CCAAT binding tran
682 26 76.5 335 2 T51106 dTDPglucose 4,6-de
683 26 76.5 335 2 T21461 hypothetical prote
684 26 76.5 336 2 T21461 homeobox protein c
685 26 76.5 337 2 T27635 conserved hypotHet
686 26 76.5 339 2 G87213

687 26 341 2 AG0195 probable exported
688 26 341 2 AG0358 conserved hypothet
689 26 343 2 S03415 hypothetical prote
690 26 344 2 A44164 secreted glycoprot
691 26 346 2 B90073 hypothetical prote
692 26 347 2 F87571 conserved hypothet
693 26 347 2 JH0657 hypothetical 37.5K
694 26 348 2 S76587 hypothetical prote
695 26 353 2 F70862 probable helix-tur
696 26 354 2 I47021 pectic enzyme secr
697 26 355 2 A45177 chemokine (C-C) re
698 26 355 2 S62692 L-amino acid oxida
699 26 356 2 B71910 integrase-recombin
700 26 358 1 WMBE38 infected cell prot
701 26 359 2 D83385 hypothetical prote
702 26 361 2 A75596 probable glucose-1
703 26 362 2 C64604 integrase/recombin
704 26 362 2 T34921 probable methyltra
705 26 362 2 T19031 hypothetical prote
706 26 363 2 S59337 RRN5 protein - yea
707 26 364 2 AI1949 hypothetical prote
708 26 369 2 JH0493 alpha-1-antichymot
709 26 370 2 D75084 carotenoid biosynt
710 26 371 2 D71063 hypothetical prote
711 26 372 2 E81350 probable RNA nucle
712 26 373 2 T32323 hypothetical prote
713 26 376 2 AH0470 UDP-N-acetylglucos
714 26 376 2 AF1978 hypothetical prote
715 26 377 2 AE1991 hypothetical prote
716 26 380 2 T11335 ubiquinol-cytochro
717 26 381 2 S12023 ubiquinol-cytochro
718 26 382 2 B61036 transforming growt
719 26 383 2 T20572 hypothetical prote
720 26 388 2 AB2372 hypothetical prote
721 26 392 2 G70594 probable MOEB- hom
722 26 393 2 D75418 homocitrate syntha
723 26 394 2 T32670 hypothetical prote
724 26 397 2 T40838 probable transport
725 26 399 2 B72649 probable adenosylh
726 26 399 2 T01345 hypothetical prote
727 26 401 2 AE1978 hypothetical prote
728 26 402 2 T12745 hypothetical prote
729 26 403 2 H90827 probable head port
730 26 403 2 F85685 probable portal pr
731 26 406 2 JS0343 cryptophan synthas
732 26 406 2 S72894 exopolyposphatase
733 26 415 2 T40338 ccaat-binding fact
734 26 417 2 S19724 kallikrein-binding
735 26 418 1 S31507 serine proteinase
736 26 418 2 S23675 contrapsin-related
737 26 418 2 JX0129 contrapsin precurs
738 26 418 2 JH0494 alpha-1-antichymot
739 26 421 2 AD2496 transposase all714
740 26 422 2 D89918 dihydrolipoamide s
741 26 423 2 AC1068 probable membrane
742 26 424 2 A35861 interferon consens
743 26 425 2 A45064 interferon consens
744 26 425 2 E71039 hypothetical prote
745 26 426 2 H95250 conserved hypothet
746 26 427 2 C83534 probable C4-dicarb
747 26 428 2 T05253 CER2 protein homol
748 26 434 2 E98115 conserved hypothet
749 26 437 2 S07040 alpha-amylase (EC
750 26 438 1 ALBH alpha-amylase (EC
751 26 439 2 F70337 hypothetical prote
752 26 441 2 T24591 hypothetical prote
753 26 443 2 H83278 hypothetical prote
754 26 445 2 T42366 L-aminoadipate-sem
755 26 448 2 AE2590 two component resp
756 26 449 2 B83456 probable transport
757 26 450 2 S57837 lymphoid-specific
758 26 455 2 T04698 hypothetical prote
759 26 456 2 F97372 probable two-compo

760 26 76.5 457 2 C82911
761 26 76.5 457 2 JC6551
762 26 76.5 460 2 AD2521
763 26 76.5 462 2 S42372
764 26 76.5 464 2 G70362
765 26 76.5 465 2 H84347
766 26 76.5 469 2 D75186
767 26 76.5 469 2 E86421
768 26 76.5 472 2 S77226
769 26 76.5 472 2 T30101
770 26 76.5 473 2 E72715
771 26 76.5 473 2 H72744
772 26 76.5 474 2 T19543
773 26 76.5 475 2 E71219
774 26 76.5 476 1 A29881
775 26 76.5 476 2 S64291
776 26 76.5 477 2 S52162
777 26 76.5 477 2 C91034
778 26 76.5 477 2 D85878
779 26 76.5 484 2 T31224
780 26 76.5 486 2 T42254
781 26 76.5 490 2 C84091
782 26 76.5 495 2 B71311
783 26 76.5 507 2 B42249
784 26 76.5 507 2 G83065
785 26 76.5 508 2 H86474
786 26 76.5 509 1 D64312
787 26 76.5 517 2 S21042
788 26 76.5 519 2 D82536
789 26 76.5 522 2 D69226
790 26 76.5 522 2 T24144
791 26 76.5 522 2 T45861
792 26 76.5 523 2 T30091
793 26 76.5 528 2 T32012
794 26 76.5 528 2 E71525
795 26 76.5 529 2 AI0986
796 26 76.5 530 2 S33673
797 26 76.5 530 2 B84130
798 26 76.5 530 2 E72129
799 26 76.5 530 2 A86494
800 26 76.5 534 2 G84713
801 26 76.5 534 2 H82244
802 26 76.5 534 2 T30019
803 26 76.5 536 1 SYECEB
804 26 76.5 536 2 E85558
805 26 76.5 536 2 A99708
806 26 76.5 542 1 A54963
807 26 76.5 546 2 F70866
808 26 76.5 547 2 T48551
809 26 76.5 550 2 AI2249
810 26 76.5 559 2 H71983
811 26 76.5 559 2 D64521
812 26 76.5 559 2 C72732
813 26 76.5 560 2 AD0359
814 26 76.5 564 2 T37934
815 26 76.5 566 2 AC0559
816 26 76.5 566 2 C90691
817 26 76.5 566 2 G85541
818 26 76.5 566 2 E64774
819 26 76.5 568 2 T38119
820 26 76.5 580 2 A72032
821 26 76.5 580 2 B86592
822 26 76.5 583 2 S56680
823 26 76.5 588 2 T04736
824 26 76.5 590 2 AI0094
825 26 76.5 593 2 T01575
826 26 76.5 597 2 C82367
827 26 76.5 606 2 G64659
828 26 76.5 607 2 S77092
829 26 76.5 610 2 S41315
830 26 76.5 612 2 T37181
831 26 76.5 614 2 G71298
832 26 76.5 622 2 JC7973

hypothetical prote
chitinase (EC 3.2.
hypothetical prote
hypothetical prote
hypothetical prote
glycerol-3-phospha
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable aspartate
probable D-lactate
hypothetical prote
hypothetical prote
ubiquinol-cytochro
hypothetical prote
sucrose hydrolase
sucrose-6 phosphat
sucrose hydrolase
traH protein homol
amino acid permeas
hypothetical prote
probable thiophene
serine-type carbox
probable malate de
hypothetical prote
conserved hypothet
cytochrome-c oxida
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable ABC trans
probable membrane
probable protein k
two-component resp
ABC transporter pr
ABC transporter pr
hypothetical prote
sensor histidine k
hypothetical prote
2,3-dihydroxybenzo
2,3-dihydroxybenzo
transcription fact
probable maltase -
fructosidase-like
exopolyposphatase
DNA primase - Heli
DNA primase - Heli
probable medium-ch
probable membrane
conserved hypothet
probable solute-bi
hypothetical prote
hypothetical prote
probable membrane
hypothetical UPF00
DNA mismatch repai
DNA mismatch repai
beta-fructofuranos
hypothetical prote
ABC-transporter tr
beta-fructofuranos
aminopeptidase P V
flagellar hook-ass
hypothetical prote
hypothetical prote
thiamin biosynthes
probable methyl-ac
synleurin - human

833 26 76.5 624 2 S74952 hypothetical prote
834 26 76.5 625 2 T40742 hypothetical integ
835 26 76.5 650 2 T15972 hypothetical prote
836 26 76.5 653 1 UZADP2 terminal protein p
837 26 76.5 653 1 UZADP5 terminal protein p
838 26 76.5 657 1 A64079 2',3'-cyclic-nucle
839 26 76.5 662 2 JC7906 sucrose 1F-fructos
840 26 76.5 664 2 T24459 hypothetical prote
841 26 76.5 670 2 T02092 beta-fructofuranos
842 26 76.5 674 2 S32230 Ca2+-transporting
843 26 76.5 681 1 H82059 2',3'-cyclic-nucle
844 26 76.5 682 1 JH0560 cyclic nucleotide-
845 26 76.5 688 2 T44343 hypothetical prote
846 26 76.5 691 2 A54741 erythrocyte membra
847 26 76.5 693 2 AI2584 hypothetical prote
848 26 76.5 697 2 H85205 potassium channel
849 26 76.5 697 2 T04931 potassium channel
850 26 76.5 698 1 JX0202 long-chain-fatty-a
851 26 76.5 702 2 C82792 methionyl-tRNA syn
852 26 76.5 718 2 C82817 ATP-dependent DNA
853 26 76.5 721 2 A39707 erythrocyte membra
854 26 76.5 728 2 H97366 ribonucleotide red
855 26 76.5 739 2 B82552 phage-related prot
856 26 76.5 739 2 F82769 phage-related tail
857 26 76.5 742 2 AE2357 hypothetical prote
858 26 76.5 743 2 B84639 probable ubiquitin
859 26 76.5 744 2 S57061 nuclear pore prote
860 26 76.5 761 2 G70393 conserved hypothet
861 26 76.5 762 2 H83415 cis/trans isomeras
862 26 76.5 764 2 A47456 down-regulated in
863 26 76.5 766 2 H83141 probable two-compo
864 26 76.5 769 1 ISBYT1 DNA topoisomerase
865 26 76.5 772 2 E82592 transcription-rela
866 26 76.5 773 2 T40694 probable rna matur
867 26 76.5 786 2 H69980 single-strand DNA-
868 26 76.5 791 2 T28815 hypothetical prote
869 26 76.5 792 2 AD0104 probable glucosida
870 26 76.5 793 2 T40285 hypothetical prote
871 26 76.5 795 1 SYECFB phenylalanine-tRNA
872 26 76.5 795 2 AE0705 phenylalanyl-tRNA
873 26 76.5 795 2 H85779 phenylalanine tRNA
874 26 76.5 795 2 D90931 phenylalanine tRNA
875 26 76.5 795 2 D82225 phenylalanyl-tRNA
876 26 76.5 795 2 AD0296 phenylalanine-tRNA
877 26 76.5 803 2 AD1282 leucyl-tRNA synthe
878 26 76.5 804 2 B89961 leucyl-tRNA synthe
879 26 76.5 804 2 B89961 leucine-tRNA ligas
880 26 76.5 804 2 D69650 leucine-tRNA ligas
881 26 76.5 805 2 T48840 transducer protein
882 26 76.5 806 2 A84060 leucyl-tRNA synthe
883 26 76.5 807 2 F85647 probable outer mem
884 26 76.5 807 2 F64844 ycds protein precu
885 26 76.5 810 2 C84845 probable outer mem
886 26 76.5 812 2 D96979 probable salt-indu
887 26 76.5 815 2 T36671 leucyl-tRNA synthe
888 26 76.5 819 2 G81698 leucine-tRNA ligas
889 26 76.5 819 2 C71544 leucyl tRNA synthe
890 26 76.5 820 2 A86510 leucine-tRNA ligas
891 26 76.5 822 2 E75523 leucine-tRNA ligas
892 26 76.5 824 2 T10020 ATP-dependent heli
893 26 76.5 824 2 F72408 leucine-tRNA ligas
894 26 76.5 825 2 JC4163 leucine-tRNA ligas
895 26 76.5 825 2 H86726 leucine-tRNA ligas
896 26 76.5 829 2 H95029 leucyl-tRNA synthe
897 26 76.5 833 2 C97901 leucine-tRNA ligas
898 26 76.5 833 2 C69200 surface proteinase
899 26 76.5 837 2 B69216 surface proteinase
900 26 76.5 840 2 C91212 probable outer mem
901 26 76.5 844 2 C86058 probable fimbrial
902 26 76.5 844 2 C87740 protein H26D21.2 [
903 26 76.5 849 2 AE3567 Arp-dependent heli
904 26 76.5 851 2 AE3567 Arp-dependent heli

26 906 76.5 853 2 AD0544 outer membrane fim
26 907 76.5 858 1 VCLJG2 env polypoteine pr
26 908 76.5 859 1 VCLJCT env polypoteine pr
26 909 76.5 867 2 G69485 DNA-directed RNA p
26 910 76.5 869 2 S77455 leucine-tRNA ligas
26 911 76.5 872 2 AD2216 leucyl-tRNA synthe
26 912 76.5 874 2 D75307 leucyl-tRNA synthe
26 913 76.5 896 2 T24169 hypothetical prote
26 914 76.5 902 2 C83964 cation-transportin
26 915 76.5 906 2 T24166 hypothetical prote
26 916 76.5 908 2 T22376 hypothetical prote
26 917 76.5 909 2 D87434 phosphoenolpyruvat
26 918 76.5 915 2 A43802 cellulase (EC 3.2.
26 919 76.5 927 1 JQ0948 A5 antigen precurs
26 920 76.5 951 2 T08987 probable cadmium-t
26 921 76.5 954 2 H71427 hypothetical prote
26 922 76.5 955 2 T48515 hypothetical prote
26 923 76.5 955 2 T00247 zinc finger protei
26 924 76.5 961 2 A47055 transposase - Erwi
26 925 76.5 967 2 T15680 hypothetical prote
26 926 76.5 968 2 C82452 probable leus prot
26 927 76.5 969 2 A70912 leucine-tRNA ligas
26 928 76.5 972 2 T10023 probable leucyl tr
26 929 76.5 973 2 A85055 hypothetical prote
26 930 76.5 980 2 T27342 probable aminoadip
26 931 76.5 981 2 T41029 env polypoteine -
26 932 76.5 985 1 VCLJSF zinc proteinase (E
26 933 76.5 989 2 T46183 probable sensor/re
26 934 76.5 992 2 A83324 leucine-tRNA ligas
26 935 76.5 994 1 SYNCLM DNA-directed DNA p
26 936 76.5 994 2 AC3480 env protein - simi
26 937 76.5 994 2 SL8739 hypothetical prote
26 938 76.5 995 2 C84145 hypothetical prote
26 939 76.5 999 2 AG2413 Ca2+-transporting
26 940 76.5 1003 2 S07526 ksr protein - frui
26 941 76.5 1003 2 T13856 conserved hypothet
26 942 76.5 1018 2 AG0703 probable oxidase l
26 943 76.5 1018 2 B90928 probable oxidase l
26 944 76.5 1018 2 F85776 conserved hypothet
26 945 76.5 1018 2 AH0293 conserved hypothet
26 946 76.5 1018 2 G64926 probable iron-sulf
26 947 76.5 1021 2 F82230 conserved hypothet
26 948 76.5 1027 2 B64187 conserved hypothet
26 949 76.5 1029 2 T00712 protein kinase hom
26 950 76.5 1039 2 S02711 cellulase (EC 3.2.
26 951 76.5 1068 1 A43322 1-phosphatidylinos
26 952 76.5 1068 1 I38110 1-phosphatidylinos
26 953 76.5 1075 2 T45570 kinesin-like prote
26 954 76.5 1075 2 D70568 hypothetical prote
26 955 76.5 1089 2 E81446 carbamoyl-phosphat
26 956 76.5 1100 2 T17270 hypothetical prote
26 957 76.5 1112 2 T02848 hsp70-related prot
26 958 76.5 1114 2 T30299 dynein heavy chain
26 959 76.5 1118 1 A49724 protein-tyrosine-p
26 960 76.5 1139 2 B70954 hypothetical prote
26 961 76.5 1172 2 S32689 parasporal crystal
26 962 76.5 1179 2 AG1463 transcription-repa
26 963 76.5 1179 2 AG1463 protein F15H18.21
26 964 76.5 1197 2 D86317 probable pol polyp
26 965 76.5 1199 2 T18348 DNA-directed DNA p
26 966 76.5 1220 1 DJBEC3 DNA-directed DNA p
26 967 76.5 1220 2 T42573 DNA-directed DNA p
26 968 76.5 1235 1 DJBEAN DNA-directed DNA p
26 969 76.5 1235 1 DJBEH7 DNA-directed DNA p
26 970 76.5 1235 1 DJBEK1 DNA-directed DNA p
26 971 76.5 1235 1 DJBEV1 DNA-directed DNA p
26 972 76.5 1240 1 DJBE21 DNA-directed DNA p
26 973 76.5 1271 2 T43269 microcystin synthe
26 974 76.5 1313 2 F96673 hypothetical prote
26 975 76.5 1313 2 F96673 mannan endo-1,4-be
26 976 76.5 1331 2 A48954 DNA-directed RNA p
26 977 76.5 1391 2 F81233 DNA-directed RNA p
26 978 76.5 1391 2 C82007 DNA-directed RNA p

979	26	76.5	1438	2	T17402	dihydroaeruginoinc
980	26	76.5	1453	2	F88640	protein F52C12.4 (
981	26	76.5	1476	1	A39901	cystic fibrosis tr
982	26	76.5	1476	1	A40303	cystic fibrosis tr
983	26	76.5	1561	2	T00248	zinc finger protei
984	26	76.5	1683	2	AF2071	WD-40 repeat prote
985	26	76.5	1691	1	A44212	genome polyprotein
986	26	76.5	1711	2	T31337	1,4-beta-glucanase
987	26	76.5	1711	2	AD1842	WD-40 repeat prote
988	26	76.5	1742	2	T17120	cellulase (EC 3.2.
989	26	76.5	1747	2	AC1842	WD-40 repeat prote
990	26	76.5	1755	2	F82618	chemotaxis-related
991	26	76.5	1808	2	T47792	hypothetical prote
992	26	76.5	1858	1	A44214	genome polyprotein
993	26	76.5	2262	2	S16664	large protein L -
994	26	76.5	2327	2	T31733	hypothetical prote
995	26	76.5	2731	1	VFIHJH	genome polyprotein
996	26	76.5	2733	2	S15760	genome polyprotein
997	26	76.5	3161	2	T30342	protein HMWP1 - Ye
998	26	76.5	4063	2	T42993	probable spectrin
999	26	76.5	4101	2	T23630	hypothetical prote
1000	26	76.5	5825	2	T12117	polyprotein - fava
ALIGNMENTS						
RESULT 1						
B86411						
protein F3M18.4 [imported] - Arabidopsis thaliana						
C;Species: Arabidopsis thaliana (mouse-ear cress)						
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004						
C;Accession: B86411						
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,						
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;						
ansen, N.F.; Hughes, B.; Huizar, L.						
Nature 408, 816-820, 2000						
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.						
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,						
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.						
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,						
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.						
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.						
A;Reference number: A86141; MUID:21016719; PMID:11130712						
A;Accession: B86411						
A;Status: preliminary						
A;Molecule type: DNA						
A;Residues: 1-486 <STO>						
A;Cross-references: UNIPROT:Q9SGQ0; UNIPARC:UPI00000ABA54; GB:AE005172; NID:g6560771; PI						
C;Genetics:						
A;Gene: F3M18.4						
A;Map position: 1						
Query Match						
Best Local Similarity 100.0%; Score 34; DB 2; Length 486;						
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	TLREWL 6				
Db	311	TLREWL 316				
RESULT 2						
S23570						
pol polyprotein homolog - fungus (Cladosporium fulvum) retrotransposon Cft-1 (fragment)						
N;Alternate names: reverse transcriptase						
C;Species: Cladosporium fulvum						
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000						
C;Accession: S23570; S19850						
R;McHale, M.T.; Roberts, I.N.; Noble, S.M.; Beaumont, C.; Whitehead, M.P.; Seth, D.; Old						
Mol. Gen. Genet. 233, 337-347, 1992						
A;Title: Cft-I: an LTR-retrotransposon in Cladosporium fulvum, a fungal pathogen of toma						
A;Reference number: S23569; MUID:92318885; PMID:1377773						

A;Accession: S23570						
A;Molecule type: DNA						
A;Residues: 1-1045 <MCH>						
A;Cross-references: UNIPARC:UPI000011E65B; EMBL:Z11866; NID:g2562; PIDN:CAA77891.1; PID:						
C;Genetics:						
A;Mobile element: retrotransposon Cft-1						
C;Keywords: polyprotein						
Query Match						
Best Local Similarity 100.0%; Score 34; DB 2; Length 1045;						
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	TLREWL 6				
Db	336	TLREWL 341				
RESULT 3						
E84853						
hypothetical protein At2g42400 [imported] - Arabidopsis thaliana						
C;Species: Arabidopsis thaliana (mouse-ear cress)						
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004						
C;Accession: E84853						
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;						
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.						
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J						
Nature 402, 761-768, 1999						
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.						
A;Reference number: A84420; MUID:20083487; PMID:10617197						
A;Accession: E84853						
A;Status: preliminary						
A;Molecule type: DNA						
A;Residues: 1-473 <STO>						
A;Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI000017A02D; GB:AE002093; NID:g4567312; PI						
C;Genetics:						
A;Gene: At2g42400						
A;Map position: 2						
Query Match						
Best Local Similarity 83.3%; Score 32; DB 2; Length 473;						
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	TLREWL 6				
		.				
Db	347	TIREWL 352				
RESULT 4						
S41808						
glucose-6-phosphate isomerase (EC 5.3.1.9), cytosolic - Arabidopsis thaliana						
N;Alternate names: phosphoglucose isomerase						
C;Species: Arabidopsis thaliana (mouse-ear cress)						
C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999						
C;Accession: S41808						
R;Thomas, B.R.; Ford, V.S.; Pichersky, E.; Gottlieb, L.D.						
Genetics 135, 895-905, 1993						
A;Title: Molecular characterization of duplicate cytosolic phosphoglucose isomerase gene						
A;Reference number: S41806; MUID:94123980; PMID:8293986						
A;Accession: S41808						
A;Molecule type: DNA						
A;Residues: 1-560 <THO>						
A;Cross-references: UNIPARC:UPI0000001752; EMBL:X69195; NID:g415922; PIDN:CAA48940.1; PI						
C;Genetics:						
A;Gene: pgIC						
A;Introns: 18/3; 44/2; 71/3; 87/3; 139/3; 172/1; 186/2; 209/1; 237/2; 253/3; 278/1; 298/;						
C;Superfamily: glucose-6-phosphate isomerase						
C;Keywords: cytosol; gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomera						
Query Match						
Best Local Similarity 83.3%; Score 32; DB 2; Length 560;						
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	TLREWL 6				

Db 227 TLREWI 232

|||||:
RESULT 5
T09154
glucose-6-phosphate isomerase (EC 5.3.1.9), cytosol - spinach
C;Species: Spinacia oleracea (spinach)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T09154
R;Nowitzki, U.; Flechner, A.; Kellermann, J.; Hasegawa, M.; Schnarrenberger, C.; Martin,
Gene 214, 205-213, 1998
A;Title: Eubacterial origin of nuclear genes for chloroplast and cytosolic glucose-6-ph
A;Reference number: Z16593; MUID:98322268; PMID:9651529
A;Accession: T09154
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-566 <NOW>
A;Cross-references: UNIPROT:O82059; UNIPARC:UPI000016DF8A; EMBL:AJ000266; NID:G3413512;
C;Genetics:
A;Gene: GPIS
C;Function:
A;Description: catalyzes reversible isomerization of glucose-6-phosphate and fructose-6-
C;Superfamily: glucose-6-phosphate isomerase
C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 94.1%; Score 32; DB 2; Length 566;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||:
Db 226 TLREWI 231

RESULT 6
S57830
glucose-6-phosphate isomerase (EC 5.3.1.9) isoenzyme 1, cytosolic - Clarkia xantiana
C;Species: Clarkia xantiana
C;Date: 28-Oct-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: S57830
R;Ford, V.S.; Thomas, B.R.; Gottlieb, L.D.
Syst. Bot. 20, 147-160, 1995
A;Title: The same duplication accounts for the PgiC genes in Clarkia xantiana and C. lew
A;Reference number: S57830
A;Accession: S57830
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-568 <FOR>
A;Cross-references: UNIPROT:P54240; UNIPARC:UPI000012AEFE; EMBL:X80666; NID:G784973; PID
C;Genetics:
A;Gene: PgiC1
A;Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C;Superfamily: glucose-6-phosphate isomerase
C;Keywords: cytosol; intramolecular oxidoreductase; isomerase

Query Match 94.1%; Score 32; DB 2; Length 568;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||:
Db 226 TLREWI 231

RESULT 7
S23542
glucose-6-phosphate isomerase (EC 5.3.1.9) 2a, cytosolic - farewell-to-spring (Clarkia 1
N;Alternate names: phosphoglucose isomerase
C;Species: Clarkia lewisii
C;Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S23542

R;Thomas, B.R.; Laudencia-Chingcuanco, D.; Gottlieb, L.D.
Plant Mol. Biol. 19, 745-757, 1992
A;Title: Molecular analysis of the plant gene encoding cytosolic phosphoglucose isomeras
A;Reference number: S23542; MUID:92353384; PMID:1643281
A;Accession: S23542
A;Molecule type: DNA
A;Residues: 1-569 <THO>
A;Cross-references: UNIPROT:P29333; UNIPARC:UPI000012AF03; EMBL:X64332; NID:G18055; PIDN
C;Genetics:
A;Gene: PgiC2a
A;Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C;Complex: homodimer
C;Superfamily: glucose-6-phosphate isomerase
C;Keywords: cytosol; gluconeogenesis; glycolysis; homodimer; intramolecular oxidoreducta

Query Match 94.1%; Score 32; DB 2; Length 569;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||:
Db 226 TLREWI 231

RESULT 8
S41806
glucose-6-phosphate isomerase (EC 5.3.1.9) 1a, cytosolic - farewell-to-spring (Clarkia 1
N;Alternate names: phosphoglucose isomerase
C;Species: Clarkia lewisii
C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S41806
R;Thomas, B.R.; Ford, V.S.; Pichersky, E.; Gottlieb, L.D.
Genetics 135, 895-905, 1993
A;Title: Molecular characterization of duplicate cytosolic phosphoglucose isomerase gene
A;Reference number: S41806; MUID:94123980; PMID:8233986
A;Accession: S41806
A;Molecule type: DNA
A;Residues: 1-569 <THO>
A;Cross-references: UNIPROT:P34796; UNIPARC:UPI000012AEFA; EMBL:X71084; NID:G415942; PII
C;Genetics:
A;Gene: PgiC1a
A;Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C;Superfamily: glucose-6-phosphate isomerase
C;Keywords: cytosol; gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomera

Query Match 94.1%; Score 32; DB 2; Length 569;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||:
Db 226 TLREWI 231

RESULT 9
S57831
glucose-6-phosphate isomerase (EC 5.3.1.9) isoenzyme 2, cytosolic - Clarkia xantiana
C;Species: Clarkia xantiana
C;Date: 28-Oct-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: S57831
R;Ford, V.S.; Thomas, B.R.; Gottlieb, L.D.
Syst. Bot. 20, 147-160, 1995
A;Title: The same duplication accounts for the PgiC genes in Clarkia xantiana and C. le
A;Reference number: S57830
A;Accession: S57831
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-569 <FOR>
A;Cross-references: UNIPROT:P54242; UNIPARC:UPI000012AF04; EMBL:X80667; NID:G784975; PII
C;Genetics:
A;Gene: PgiC2
A;Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C;Superfamily: glucose-6-phosphate isomerase

A;Molecule type: DNA
A;Residues: 1-158 <BLAT>
A;Cross-references: UNIPROT:P75815; UNIPARC:UPI00000480CD; GB:AE000187; GB:U00096; NID:9
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybjN
C;Superfamily: Escherichia coli hypothetical protein ybjN

Query Match 91.2%; Score 31; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 11 TLRQWL 16

RESULT 15
E90745
probable sensory transduction regulator [imported] - Escherichia coli (strain O157:H7, s
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: E90745
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E90745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-158 <HAY>
A;Cross-references: UNIPROT:P75815; UNIPARC:UPI00000480CD; GB:BA000007; PIDN:BAB34356.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs0933
C;Superfamily: Escherichia coli hypothetical protein ybjN

Query Match 91.2%; Score 31; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 11 TLRQWL 16

RESULT 16
C83872
hypothetical protein BH1779 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83872
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83872
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-287 <STO>
A;Cross-references: UNIPROT:Q9KBZ5; UNIPARC:UPI00000C3CC4; GB:AP001513; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1779
C;Superfamily: scyllo-inosamine-4-phosphate amidinotransferase

Query Match 91.2%; Score 31; DB 2; Length 287;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|

Db 114 TLKEWL 119

RESULT 17
B82535
conserved hypothetical protein XF2609 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: B82535
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82535
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <SIM>
A;Cross-references: UNIPARC:UPI0000165A87; GB:AE004068; GB:AE003849; NID:g9107832; PIDN:
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2609

Query Match 91.2%; Score 31; DB 2; Length 299;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 158 TLRDWL 163

RESULT 18
S34816
nitrogenase cofactor synthesis protein nifs - Rhodobacter capsulatus
N;Contains: L-cysteine sulfurtransferase (EC 2.8.1.-)
C;Species: Rhodobacter capsulatus
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S34816
R;Masepohl, B.; Angermueller, S.; Hennecke, S.; Huebner, P.; Moreno-Vivian, C.; Klipp, V
Mol. Gen. Genet. 238, 369-382, 1993
A;Title: Nucleotide sequence and genetic analysis of the Rhodobacter capsulatus ORF6-nit
A;Reference number: S34814; MUID:93261420; PMID:8492805
A;Accession: S34816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <MAS>
A;Cross-references: UNIPROT:Q07177; UNIPARC:UPI0000130115; EMBL:X68444; NID:g297915; PIR
C;Genetics:
A;Gene: nifs
C;Superfamily: nitrogen fixation protein nifs
C;Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase
F;203/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
F;325/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 384;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 267 TLRDWL 272

RESULT 19
S01825
transforming growth factor beta-3 precursor - pig
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C/Accession: S01825
R/Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mas
EMBO J. 7, 3737-3743, 1988
A/Title: A new type of transforming growth factor-beta, TGF-beta3.
A/Reference number: S01824; MUID:89091120; PMID:3208746
A/Accession: S01825
A/Molecule type: mRNA
A/Residues: 1-409 <DER>
A/Cross-references: UNIPROT:P15203; UNIPARC:UPI0000136CB9; EMBL:X14150; NID:g2127; PIDN:
C/Superfamily: inhibin
C/Keywords: growth factor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-297/Domain: propeptide #status predicted <PRO>
F;298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 91.2%; Score 31; DB 2; Length 409;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 205 TVREWL 210

RESULT 20
A41397
transforming growth factor beta-3 precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
C/Accession: A41397; A61039; A61225
R/Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
Mol. Endocrinol. 3, 1926-1934, 1989
A/Title: Complementary DNA cloning of the murine transforming growth factor-beta3 (TGFbe
t tissues.
A/Reference number: A41397; MUID:90190650; PMID:2628730
A/Accession: A41397
A/Molecule type: mRNA
A/Residues: 1-410 <MIL>
A/Cross-references: UNIPROT:PI7125; UNIPARC:UPI0000020D6C; GB:M32745; NID:g201949; PIDN:
R/Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.
Growth Factors 3, 139-146, 1990
A/Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-beta3.
A/Reference number: A61039; MUID:91000714; PMID:2206556
A/Accession: A61039
A/Molecule type: mRNA
A/Residues: 1-410 <DEN>
A/Cross-references: UNIPARC:UPI0000020D6C
R/Watrin, F.; Scotto, L.; Assolian, R.K.; Wolgemuth, D.J.
Cell Growth Differ. 2, 77-83, 1991
A/Title: Cell lineage specificity of expression of the murine transforming growth factor
A/Reference number: A61225; MUID:91299576; PMID:2069871
A/Accession: A61225
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 285-410 <WAT>
A/Cross-references: UNIPARC:UPI0000176557
C/Superfamily: inhibin
C/Keywords: glycoprotein; growth factor; growth regulation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-298/Domain: propeptide #status predicted <PRO>
F;259-261/Region: cell attachment (R-G-D) motif
F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>
F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 205 TVREWL 210

RESULT 21
A55706
transforming growth factor beta-3 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 09-Jul-2004
C/Accession: A55706; B40699; S36042
R/Wang, J.; Kuliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Tseu, I.; Post, M.
J. Biol. Chem. 270, 2722-2728, 1995
A/Title: Cloning and expression of glucocorticoid-induced genes in fetal rat lung fibrobl
A/Reference number: A55706; MUID:95155340; PMID:7852342
A/Accession: A55706
A/Molecule type: mRNA
A/Residues: 1-410 <WAN>
A/Cross-references: UNIPROT:Q07258; UNIPARC:UPI000017655A; GB:U03491
A/Note: it is uncertain whether Met-1 is the initiator
R/McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.
J. Cell Biol. 121, 1397-1407, 1993
A/Title: A role for TGF-beta in oligodendrocyte differentiation.
A/Reference number: A40699; MUID:93286190; PMID:8509457
A/Accession: B40699
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 157-211 <MCK>
A/Cross-references: UNIPARC:UPI0000170B3A; EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PII
C/Superfamily: inhibin

Query Match 91.2%; Score 31; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 205 TVREWL 210

RESULT 22
A34939
transforming growth factor beta-3 precursor - chicken
C/Species: Gallus gallus (chicken)
C/Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C/Accession: A34939; S25850; S36125; S36124; I51181
R/Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.
Mol. Endocrinol. 2, 747-755, 1988
A/Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth fact
A/Reference number: A34939; MUID:89096966; PMID:3211158
A/Accession: A34939
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-412 <JAK>
A/Cross-references: UNIPROT:P16047; UNIPARC:UPI00001713BB; GB:M31154; NID:g212758; PIDN:;
R/Burt, D.W.; Paton, I.R.; Dey, B.R.
J. Mol. Endocrinol. 7, 175-183, 1991
A/Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and
A/Reference number: S25850; MUID:92134496; PMID:1840616
A/Accession: S25850
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-117 <BUR>
A/Cross-references: UNIPARC:UPI000017146B; EMBL:X58127; NID:g63815; PIDN:CAA41128.1; PID
A/Accession: S36125
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 119-172 <BU2>

RESULT 25
T19971
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19971
R;Burton, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19206
A;Accession: T19971
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-505 <WIL>
A;Cross-references: UNIPROT:Q93383; UNIPARC:UPI0000179966; EMBL:Z81449; PIDN:CAB03762.1;
A;Experimental source: clone C46F11
C;Genetics:
A;Gene: CESP:C46F11.5a
A;Map position: 3
A;Introns: 38/2; 89/3; 119/1; 178/1; 195/1; 260/3; 313/2; 366/3; 407/1; 450/1; 480/3

Query Match 91.2%; Score 31; DB 2; Length 505;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 252 TLKEWL 257

RESULT 26
T19973
hypothetical protein C46F11.5b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19973
R;Burton, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19206
A;Accession: T19973
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-506 <WIL>
A;Cross-references: UNIPROT:O62100; UNIPARC:UPI0000179965; EMBL:Z81449; PIDN:CAB03764.1;
A;Experimental source: clone C46F11
C;Genetics:
A;Gene: CESP:C46F11.5b
A;Map position: 3
A;Introns: 38/2; 89/3; 119/1; 178/1; 195/1; 260/3; 313/2; 366/3; 407/1; 450/1; 481/3

Query Match 91.2%; Score 31; DB 2; Length 506;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 252 TLKEWL 257

RESULT 27
A96018
probable regulatory protein, possibly two-component response regulator SMB20702 [imported]
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A96018
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: A96018
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-554 <KUR>

A;Cross-references: UNIPROT:Q92TT6; UNIPARC:UPI00000CB867; GB:AL591985; PIDN:CAC49809.1;
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB20702
A;Genome: plasmid

Query Match 91.2%; Score 31; DB 2; Length 554;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 527 TLRDWL 532

RESULT 28
A56126
peroxisomal targeting signal 1 receptor - human
N;Alternate names: peroxisomal C-terminal targeting signal 1 import receptor; peroxisomal
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 31-Dec-2004
C;Accession: A56126; I38142; I38925; S53100
R;Fransen, M.; Brees, C.; Baumgart, E.; Vanhooren, J.C.T.; Baes, M.; Mannaerts, G.P.; Var
J. Biol. Chem. 270, 7731-7736, 1995
A;Title: Identification and characterization of the putative human peroxisomal C-terminal
A;Reference number: A56126; MUID:95221441; PMID:7706321
A;Accession: A56126
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-639 <FRA>
A;Cross-references: UNIPROT:P50542; UNIPARC:UPI000016AE82; GB:X84899; NID:g6955565; PIDN:C
R;Wiemer, E.A.; Nuttley, W.M.; Bertolaet, B.L.; Li, X.; Francke, U.; Wheelock, M.J.; Anne
J. Cell Biol. 130, 51-65, 1995
A;Title: Human peroxisomal targeting signal-1 receptor restores peroxisomal protein import
A;Reference number: A56735; MUID:95310365; PMID:7790377
A;Accession: I38142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-214,252-639 <WIE>
A;Cross-references: UNIPARC:UPI00001316E7; EMBL:Z48054; NID:g732797; PIDN:CAA88131.1; PII
R;Dodt, G.; Braverman, N.; Wong, C.; Moser, A.; Moser, H.W.; Watkins, P.; Valle, D.; Goul
Nature Genet. 9, 115-125, 1995
A;Title: Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the
A;Reference number: I38925; MUID:95233555; PMID:7719337
A;Accession: I38925
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-214,252-424,'1',426-639 <DOD>
A;Cross-references: UNIPARC:UPI000016A16C; EMBL:U19721; NID:g694004; PIDN:AAC50103.1; PII
C;Genetics:
A;Gene: GDB:PXR1; PEX5; PTS1R
A;Cross-references: GDB:433739; OMIM:600414
A;Map position: l2p13-l2p13
C;Function:
A;Description: binds the carboxyl-terminal tripeptide Ser-[Lys/ARG/HIS]-Leu, SKL, or per
C;Keywords: peroxisome biogenesis; protein transport
F;452-485/Domain: tetratricopeptide repeat homology <TT1>
F;488-521/Domain: tetratricopeptide repeat homology <TT2>
F;522-555/Domain: tetratricopeptide repeat homology <TT3>

Query Match 91.2%; Score 31; DB 2; Length 639;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 425 TLRDWL 430

RESULT 29
E96508
hypothetical protein T12C22.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96508
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-936 <STO>
A;Cross-references: UNIPROT:Q9LPD9; UNIPARC:UPI00000AA9E4; GB:AE005173; NID:g8656002; PI
C;Genetics:
A;Gene: T12C22.19
A;Map position: 1
C;Superfamily: replication licensing factor MCM2; MCM homology

Query Match 91.2%; Score 31; DB 2; Length 936;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 204 TLREWL 209

RESULT 30
T48216
hypothetical protein T20L15.220 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 31-Dec-2004
C;Accession: T48216
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: 224488
A;Accession: T48216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-984 <BEV>
A;Cross-references: UNIPROT:Q9LZV1; UNIPARC:UPI0000048AAF; EMBL:AL162351
A;Experimental source: cultivar Columbia; BAC clone T20L15
C;Genetics:
A;Map position: 5
A;Introns: 29/1; 164/2; 212/2; 236/2; 260/2; 285/2; 308/2; 331/2; 355/2; 384/2; 407/2; 5
A;Note: T20L15.220
C;Superfamily: Receptor-like protein kinase

Query Match 91.2%; Score 31; DB 2; Length 984;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 739 TLRDWL 744

RESULT 31
F70697

probable arabinosyltransferase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70697
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70697
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1094 <COL>
A;Cross-references: UNIPROT:P72060; UNIPARC:UPI0000129E94; GB:Z80343; GB:AL123456; NID:g
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: embA
C;Superfamily: probable arabinosyl transferase

Query Match 91.2%; Score 31; DB 2; Length 1094;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 232 TLRDWL 237

RESULT 32
T00390
KIAA0614 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00390; T17254
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00390
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1630 <ISH>
A;Cross-references: UNIPROT:Q9UFT6; UNIPARC:UPI000003B446; EMBL:AB014514; NID:g3327041;
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17254
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1515-1630 <KOE>
A;Cross-references: UNIPARC:UPI0000073F07; EMBL:AL117469
A;Experimental source: adult uterus; clone DKFZp586O1022
C;Genetics:
A;Note: KIAA0614; DKFZp586O1022.1

Query Match 91.2%; Score 31; DB 2; Length 1630;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||:|
Db 165 TVREWL 170

RESULT 33
H83343
probable non-ribosomal peptide synthetase PA2424 [imported] - Pseudomonas aeruginosa (s
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83343
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83343
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4342 <STO>
A;Cross-references: UNIPROT:Q9I157; UNIPARC:UPI0000110218; GB:AE004669; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2424
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;60-553/Domain: acetate-CoA ligase homology <ACL1>
F;584-652/Domain: acyl carrier protein homology <ACP1>
F;1174-1622/Domain: acetate-CoA ligase homology <ACL2>
F;1637-1705/Domain: acyl carrier protein homology <ACP2>
F;2232-2689/Domain: acetate-CoA ligase homology <ACL3>
F;2706-2773/Domain: acyl carrier protein homology <ACP3>
F;3759-4230/Domain: acetate-CoA ligase homology <ACL4>
F;4248-4316/Domain: acyl carrier protein homology <ACP4>
F;1669,2738,4280/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 4342;
Best Local Similarity 83.3%; Pred. No. 2.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
Db 3560 TVREWL 3565

RESULT 34
S10865
early E4 13K protein - human adenovirus 12
C;Species: Mastadenovirus h12 (human adenovirus 12)
C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S10865; S33953
R;Hogenkamp, T.; Esche, H.
Nucleic Acids Res. 18, 3065-3066, 1990
A;Title: Nucleotide sequence of the right 10% of adenovirus type 12 DNA encoding the en
A;Reference number: S10860; MUID:90272430; PMID:2349112
A;Accession: S10865
A;Molecule type: DNA
A;Residues: 1-120 <HOG>
A;Cross-references: UNIPROT:P36709; UNIPARC:UPI0000036E4C; EMBL:X51800
A;Note: the authors translated the codon TGG for residue 83 as Tyr
R;Sprengel, J.
submitted to the EMBL Data Library, June 1993
A;Reference number: S33928
A;Accession: S33953
A;Molecule type: DNA
A;Residues: 1-120 <SPR>
A;Cross-references: UNIPARC:UPI0000036E4C; EMBL:X73487; NID:G313361; PIDN:CAA51902.1; PI
C;Superfamily: adenovirus early E4 13K protein

Query Match 88.2%; Score 30; DB 2; Length 120;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
:|||||
Db 53 SLREWL 58

RESULT 35
D72305
hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: D72305
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-158 <ARN>
A;Cross-references: UNIPROT:Q9X0A5; UNIPARC:UPI00000C1322; GB:AE001763; GB:AE000512; NID:
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1012

Query Match 88.2%; Score 30; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
Db 148 TLRKWL 153

RESULT 36
G89799
hypothetical protein SA0330 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 31-Dec-2004
C;Accession: G89799
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: G89799
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-181 <KUR>
A;Cross-references: UNIPROT:Q99WN5; UNIPARC:UPI00000CACCA; GB:BA000018; PID:gl3700256; P
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA0330
C;Superfamily: ribosomal protein serine N-acetyltransferase

Query Match 88.2%; Score 30; DB 2; Length 181;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
:|||||
Db 33 SLREWL 38

RESULT 37
S52461
hypothetical protein 233 - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S52461
R;Willems, H.; Thiele, D.; Valkova, D.
submitted to the EMBL Data Library, February 1995
A;Reference number: S52460
A;Accession: S52461
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <WIL>
A;Cross-references: UNIPROT:Q45912; UNIPARC:UPI00000B946E; EMBL:X84722; NID:G682754; PID:

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 38
S38228
hypothetical protein - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S38228
R;Thiele, D.; Willems, H.; Haas, M.; Krauss, H.
submitted to the EMBL Data Library, October 1993
A;Reference number: S38215
A;Accession: S38228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <THI>
A;Cross-references: UNIPROT:Q45935; UNIPARC:UPI00000BBE8C; EMBL:X75356; NID:g407370; PID

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 39
S53104
hypothetical protein 233 - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53104
R;Ritter, M.; Thiele, D.; Willems, H.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53104
A;Accession: S53104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <RIT>
A;Cross-references: UNIPROT:Q45912; UNIPROT:Q45935; UNIPROT:Q52877; UNIPARC:UPI000017970

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 40
E69189
hypothetical protein MTH672 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: E69189
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: E69189
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-239 <MTH>
A;Cross-references: UNIPROT:Q26768; UNIPARC:UPI0000062B47; GB:AE000846; GB:AE000666; NID
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH672

A;Start codon: GTG

Query Match 88.2%; Score 30; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREW 5
|||||
Db 91 TLREW 95

RESULT 41
A30227
hypothetical protein 2BE2121 - Chinese hamster (fragment)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 01-Dec-2000
C;Accession: A30227; B30227
R;Foreman, P.K.; Hamlin, J.L.
Mol. Cell. Biol. 9, 1137-1147, 1989
A;Title: Identification and characterization of a gene that is coamplified with dihydrof
A;Reference number: A30227; MUID:89261717; PMID:2725490
A;Accession: A30227
A;Molecule type: mRNA
A;Residues: 1-263 <FOR>
A;Cross-references: UNIPARC:UPI00000E6589; GB:M23159; NID:g191006; PIDN:AAA36962.1; PID:
A;Note: the source is Chinese hamster ovary cells
A;Accession: B30227
A;Molecule type: mRNA
A;Residues: 215-263 <FO2>
A;Cross-references: UNIPARC:UPI0000179949; GB:M23160; NID:g191008
A;Note: this ORF is not annotated in GenBank entry CIUDHFRCA, release 109.0
A;Note: the source is Chinese hamster ovary cells

Query Match 88.2%; Score 30; DB 2; Length 263;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 76 TLQEWL 81

RESULT 42
F83042
hypothetical protein PA4834 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83042
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83042
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-284 <STO>
A;Cross-references: UNIPROT:Q9HUX6; UNIPARC:UPI00000C5E2F; GB:AE004896; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4834

Query Match 88.2%; Score 30; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREW 5
|||||
Db 204 TLREW 208

RESULT 43

F86431
hypothetical protein TS18.8 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86431
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: F86431
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-292 <STO>
A/Cross-references: UNIPROT:Q9SA78; UNIPARC:UPI000009F92E; GB:AB005172; NID:g4587519; PI
C/Genetics:
A/Map position: 1

Query Match 88.2%; Score 30; DB 2; Length 292;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
:|||||
Db 93 SLREWL 98

RESULT 44
S77938
EBNA-LP protein - human herpesvirus 4
C/Species: human herpesvirus 4, Epstein-Barr virus
C/Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 05-Oct-2004
C/Accession: S77938; S42441
R/Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
submitted to the EMBL Data Library, September 1987
A/Reference number: S77938
A/Accession: S77938
A/Molecule type: mRNA
A/Residues: 1-308 <SAM>
A/Cross-references: UNIPROT:Q69136; UNIPARC:UPI00000F09AD; EMBL:M13940; NID:g330401; PID
R/Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
Proc. Natl. Acad. Sci. U.S.A. 83, 5096-5100, 1986
A/Title: Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a p
A/Reference number: S42440; MUID:86259739; PMID:3460083
A/Accession: S42441
A/Molecule type: mRNA
A/Residues: 'M', 134-308 <SAM>
A/Cross-references: UNIPARC:UPI0000177CC9; EMBL:M13940
C/Keywords: alternative splicing

Query Match 88.2%; Score 30; DB 2; Length 308;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
:|||||
Db 275 SLREWL 280

RESULT 45
F82601
5'-nucleotidase XF2089 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: F82601
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: F82601
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-319 <SIM>
A/Cross-references: UNIPROT:Q9PBQ1; UNIPARC:UPI00000C2928; GB:AE004024; GB:AE003849; NID
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tauhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF2089

Query Match 88.2%; Score 30; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 5
:|||||
Db 257 TLREWL 261

RESULT 46
A29711
deacetoxycephalosporin C synthetase - fungus (Acremonium sp.)
N/Alternate names: expandase; hydroxylase
C/Species: Acremonium sp.
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 31-Dec-2004
C/Accession: A29711; A41864
R/Samson, S.M.; Dotzlaef, J.E.; Slisz, M.L.; Becker, G.W.; Van Frank, R.M.; Veal, L.E.; Ye
Bio/Technology 5, 1207-1214, 1987
A/Title: Cloning and expression of the fungal expandase/hydroxylase gene involved in cep
A/Reference number: A29711
A/Accession: A29711
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-332 <SAM>
A/Cross-references: UNIPROT:P11935; UNIPARC:UPI0000044675
A/Note: the source is designated as Cephalosporium acremonium
R/Gutierrez, S.; Velasco, J.; Fernandez, F.J.; Martin, J.F.
J. Bacteriol. 174, 3056-3064, 1992
A/Title: The cefG gene of Cephalosporium acremonium is linked to the cefEF gene and enco
A/Reference number: A41864; MUID:92234966; PMID:1569032
A/Accession: A41864
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-76,'L', <GUT>
A/Cross-references: UNIPARC:UPI0000175256
A/Experimental source: strain C10
A/Note: sequence extracted from NCBI backbone (NCBIN:104773, NCBIP:97574); this ORF is n
A/Note: the source is designated as Cephalosporium acremonium
C/Superfamily: isopenicillin N synthase

Query Match 88.2%; Score 30; DB 2; Length 332;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TLREWL 6
:|||||
Db 294 TPREWL 299

RESULT 47

E90998
hypothetical protein ECs2957 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: E90998
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E90998
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-338 <HAY>
A;Cross-references: UNIPROT:Q9EYD4; UNIPARC:UPI00000B38DF; GB:BA000007; PIDN:BAB36380.1;
A;Experimental source: strain O157:H7, substrain R1MD 0509952
C;Genetics:
A;Gene: ECs2957

Query Match 88.2%; Score 30; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 5
|||||
Db 20 TLREW 24

RESULT 48

D85818
unknown protein encoded within prophage CP-933U [imported] - Escherichia coli (strain O1
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85818
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85818
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <STO>
A;Cross-references: UNIPROT:Q8X8Z2; UNIPARC:UPI000016582A; GB:AE005174; NID:g12516109; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3092

Query Match 88.2%; Score 30; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 5
|||||
Db 28 TLREW 32

RESULT 49

AF2963
phosphomannose isomerase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2963
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2963

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-394 <KUR>

A;Cross-references: UNIPROT:Q8UAR2; UNIPARC:UPI00000D2161; GB:AE008689; PIDN:AAL44124.1;

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3311

A;Map position: linear chromosome

Query Match 88.2%; Score 30; DB 2; Length 394;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 6
|||||
Db 18 TLRKWL 23

RESULT 50

G98319
mannose-6-phosphate isomerase (phosphomannose isomerase) (pmi) (phosphohexomutase) [limpo
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: G98319
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: G98319
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <KUR>
A;Cross-references: UNIPROT:Q8UAR2; UNIPARC:UPI00000D2161; GB:AE007870; PIDN:AAK90081.1;
C;Genetics:
A;Gene: AGR_L_3017
A;Map position: linear chromosome

Query Match 88.2%; Score 30; DB 2; Length 394;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 6
|||||
Db 18 TLRKWL 23

Search completed: May 12, 2006, 10:52:12

Job time : 15.3821 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 42.4615 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-34
Perfect score: 34
Sequence: 1 TLREWL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	100.0	78	2	Q8ZZ85_PYRAE	Q8zz85 pyrobaculum
2	34	100.0	78	2	Q50WB5_ENTHI	Q50wb5 entamoeba h
3	34	100.0	190	2	Q9SHP3_ARATH	Q9shp3 arabidopsis
4	34	100.0	213	2	Q4T3D3_TETNG	Q4t3d3 tetraodon n
5	34	100.0	243	2	Q7TNX8_RAT	Q7tnx8 rattus norv
6	34	100.0	258	1	LRP16_RAT	Q8k4g6 rattus norv
7	34	100.0	265	1	KSGA_AZOSE	Q5p7j1 azoarcus sp
8	34	100.0	277	2	Q87IG5_VIBPA	Q87ig5 vibrio para
9	34	100.0	296	2	Q9KX21_OLICA	Q9kx21 oligotropha
10	34	100.0	325	1	LRP16_HUMAN	Q9bq69 homo sapien
11	34	100.0	391	2	Q75PR0_PHYPA	Q75pr0 physcomitre
12	34	100.0	473	2	Q7PSJ0_ANOGA	Q7psj0 anopheles g
13	34	100.0	486	2	Q9SGQ0_ARATH	Q9sgq0 arabidopsis
14	34	100.0	522	2	Q5BGN6_EMENI	Q5bgn6 aspergillus
15	34	100.0	556	2	Q6XPC4_PHYIN	Q6xpc4 phytophthor
16	34	100.0	557	2	Q4HCY6_9DEIO	Q4hcy6 deinococcus
17	34	100.0	597	2	Q5M7M1_XENTR	Q5m7m1 xenopus tro
18	34	100.0	637	2	Q6FLE9_CANGA	Q6fle9 candida gla
19	34	100.0	963	2	Q9VYJ0_DROME	Q9vyj0 drosophila
20	34	100.0	1243	2	Q9UVC1_CLAFU	Q9uvc1 cladosporiu
21	34	100.0	2066	2	Q9Z3T8_PSESX	Q9z3t8 pseudomonas
22	34	100.0	2066	2	Q87W69_PSESM	Q87w69 pseudomonas
23	34	100.0	5072	2	Q6JD66_TOXGO	Q6jd66 toxoplasma
24	32	94.1	81	2	Q6FEL5_ACIAD	Q6fel5 acinetobact
25	32	94.1	134	2	Q70SL4_9MYRT	Q70sl4 hauya heyde
26	32	94.1	134	2	Q70SM0_9MYRT	Q70sm0 circaea cor
27	32	94.1	144	2	Q70SL7_9MYRT	Q70sl7 fuchsia cyl
28	32	94.1	176	2	Q656A5_ORYSA	Q656a5 oryza sativ
29	32	94.1	195	2	Q8D9L1_VIBVU	Q8d9l1 vibrio vuln
30	32	94.1	209	2	Q6N1X5_RHOPA	Q6n1x5 rhodopseudo
31	32	94.1	269	2	Q70SKI_9MYRT	Q70sk1 lopezia gra

32	94.1	281	2	Q70SK3_9MYRT	Q70sk3 megacorax g
32	94.1	307	2	Q8RW74_CLAEP	Q8rw74 clarkia epi
34	94.1	313	1	TOC34_ARATH	Q38906 arabidopsis
35	94.1	328	2	Q5V5F2_HALMA	Q5v5f2 haloarcula
36	94.1	342	2	Q6MB42_PARUW	Q6mb42 parachlamyd
37	94.1	357	2	Q22092_9LILI	Q22092 dioscorea n
38	94.1	357	2	Q22093_9LILI	Q22093 dioscorea q
39	94.1	357	2	Q22094_9LILI	Q22094 dioscorea s
40	94.1	357	2	Q22095_9LILI	Q22095 dioscorea t
41	94.1	357	2	Q24644_9LILI	Q24644 dioscorea g
42	94.1	368	2	Q4SIN4_TETNG	Q4sin4 tetraodon n
43	94.1	371	2	Q70SK2_9MYRT	Q70sk2 hauya elega
44	94.1	371	2	Q70SK7_9MYRT	Q70sk7 epilobium c
45	94.1	373	2	Q70SK5_9MYRT	Q70sk5 epilobium b
46	94.1	380	2	Q8L6V2_9MYRT	Q8l6v2 clarkia dud
47	94.1	380	2	Q8L6V3_9MYRT	Q8l6v3 clarkia dud
48	94.1	380	2	Q8L6V4_9MYRT	Q8l6v4 clarkia dud
49	94.1	380	2	Q8RW61_CLAUN	Q8rw61 clarkia ung
50	94.1	381	2	Q70SL1_LUDPE	Q70sl1 ludwigia pe
51	94.1	402	2	Q6DIP8_XENTR	Q6dip8 xenopus tro
52	94.1	411	2	Q23903_9LILI	Q23903 dioscorea t
53	94.1	429	2	Q8RW83_9MYRT	Q8rw83 clarkia del
54	94.1	436	2	Q22078_9LILI	Q22078 dioscorea q
55	94.1	443	2	Q9P858_PHANO	Q9p858 phaecosphaer
56	94.1	450	2	Q9SLB9_ARATH	Q9slb9 arabidopsis
57	94.1	476	2	Q8VXR4_9MYRT	Q8vxxr4 clarkia del
58	94.1	491	2	Q70SK9_GAULI	Q70sk9 gaura lindh
59	94.1	499	2	Q8L6U8_9MYRT	Q8l6u8 clarkia lin
60	94.1	499	2	Q8L6U9_9MYRT	Q8l6u9 clarkia lin
61	94.1	501	2	Q70SL0_9MYRT	Q70sl0 calylophus
62	94.1	501	2	Q8RW62_CLAUN	Q8rw62 clarkia ung
63	94.1	501	2	Q8RW72_9MYRT	Q8rw72 clarkia mod
64	94.1	516	2	Q8RW64_9MYRT	Q8rw64 clarkia sim
65	94.1	526	2	Q8RW65_9MYRT	Q8rw65 clarkia sim
66	94.1	526	2	Q8RW66_9MYRT	Q8rw66 clarkia sim
67	94.1	526	2	Q8RW79_9MYRT	Q8rw79 clarkia del
68	94.1	527	2	Q8RUS8_9MYRT	Q8rus8 clarkia sim
69	94.1	527	2	Q8RW63_9MYRT	Q8rw63 clarkia sim
70	94.1	527	2	Q8RW76_9MYRT	Q8rw76 clarkia del
71	94.1	527	2	Q8RW81_9MYRT	Q8rw81 clarkia del
72	94.1	529	2	Q839B8_ENTFA	Q839b8 enterococcu
73	94.1	534	2	Q88Z37_LACPL	Q88z37 lactobacill
74	94.1	559	2	Q8L6V0_9MYRT	Q8l6v0 clarkia het
75	94.1	560	1	G6PI_ARALP	Q9fxm5 arabis gemm
76	94.1	560	1	G6PI_ARALP	Q9fxm4 arabidopsis
77	94.1	560	1	G6PI_ARATH	P34795 arabidopsis
78	94.1	560	2	Q7G121_9BRAS	Q7g121 leavenworth
79	94.1	560	2	Q7X833_ARAGE	Q7x833 arabis gemm
80	94.1	560	2	Q7X8A0_ARAGE	Q7x8a0 arabis gemm
81	94.1	560	2	Q7X9W9_ARAGE	Q7x9w9 arabis gemm
82	94.1	560	2	Q7X9X0_ARAGE	Q7x9x0 arabis gemm
83	94.1	560	2	Q7X9X1_ARAGE	Q7x9x1 arabis gemm
84	94.1	560	2	Q7X9X2_ARAGE	Q7x9x2 arabis gemm
85	94.1	560	2	Q7X9X3_ARAGE	Q7x9x3 arabis gemm
86	94.1	560	2	Q7X9X4_ARAGE	Q7x9x4 arabis gemm
87	94.1	560	2	Q7X9X5_ARAGE	Q7x9x5 arabis gemm
88	94.1	560	2	Q7X9X6_ARAGE	Q7x9x6 arabis gemm
89	94.1	560	2	Q8H9A5_9BRAS	Q8h9a5 crucihimala
90	94.1	560	2	Q8H9A6_ARAGL	Q8h9a6 arabis glab
91	94.1	560	2	Q94JT1_ARATH	Q94jt1 arabidopsis
92	94.1	560	2	Q546J1_ARATH	Q546j1 arabidopsis
93	94.1	566	1	G6PI_SPIOL	Q82059 spinacia ol
94	94.1	567	2	Q23904_9LILI	Q23904 dioscorea t
95	94.1	567	2	Q23905_9LILI	Q23905 dioscorea t
96	94.1	567	2	Q23906_9LILI	Q23906 dioscorea t
97	94.1	567	2	Q23907_9LILI	Q23907 dioscorea t
98	94.1	567	2	Q24604_9LILI	Q24604 dioscorea t
99	94.1	567	2	Q24615_9LILI	Q24615 dioscorea t
100	94.1	567	2	Q7DN41_9LILI	Q7dn41 dioscorea t
101	94.1	567	2	Q7DN41_9LILI	Q7dn41 dioscorea t
102	94.1	568	1	G6PI1_CLAAR	P54234 clarkia arc
103	94.1	568	1	G6PI1_CLAAR	P54236 clarkia fra
104	94.1	568	1	G6PI1_CLAMI	P54237 clarkia mil

105	32	94.1	568	1	G6PI1_CLAWI	P54239	clarkia wil
106	32	94.1	568	1	G6PI1_CLAXA	P54240	clarkia xan
107	32	94.1	568	1	G6PI_OENME	P54243	oenothera m
108	32	94.1	568	2	Q9SMJ2_9MYRT	P54235	clarkia gra
109	32	94.1	569	1	G6PI1_CLACO	P34796	clarkia lew
110	32	94.1	569	1	G6PI1_CLALE	P54241	clarkia con
111	32	94.1	569	1	G6PI2_CLACO	P29333	clarkia lew
112	32	94.1	569	1	G6PI2_CLALE	P54242	clarkia xan
113	32	94.1	569	1	G6PI2_CLAXA	Q8rw7l	clarkia mod
114	32	94.1	569	2	Q8RW7l_9MYRT	Q9smj1	clarkia gra
115	32	94.1	569	2	Q9SMJ1_9MYRT	Q8vxr1	clarkia epi
116	32	94.1	569	2	Q8VXR1_CLAEP	Q8vxr2	clarkia epi
117	32	94.1	569	2	Q8VXR2_CLAEP	P54238	clarkia ros
118	32	94.1	570	1	G6PI1_CLARO	Q5z3n5	nocardia fa
119	32	94.1	614	2	Q5Z3N5_NOCFA	Q8drt3	streptococc
120	32	94.1	834	2	Q8DRT3_STRMU	Q4ggf6	leishmania
121	32	94.1	971	2	Q4QGF6_LEIMA	Q8itl2	trypanosoma
122	32	94.1	1142	2	Q8ITL2_9TRYP	Q57w38	trypanosoma
123	32	94.1	1142	2	Q57W38_9TRYP	Q5bul3	trypanosoma
124	32	94.1	1142	2	Q5BU13_9TRYP	Q4hdk1	campylobact
125	31	91.2	70	2	Q4HDK1_CAMCO	Q9rs35	deinococcus
126	31	91.2	85	2	Q9RS35_DEIRA	Q8frd0	corynebacte
127	31	91.2	99	2	Q8FRD0_COREF	Q10i39	schizosacch
128	31	91.2	105	1	YAS3_SCHPO	Q8dhx7	synechococc
129	31	91.2	129	2	Q8DHX7_SYNEL	P758i5	escherichia
130	31	91.2	158	1	YBJN_ECOLI	Q57r86	salmonella
131	31	91.2	158	2	Q57R86_SALCH	Q5pgm6	salmonella
132	31	91.2	158	2	Q5PGM6_SALPA	Q8z847	salmonella
133	31	91.2	158	2	Q8Z847_SALTI	Q8zqk1	salmonella
134	31	91.2	158	2	Q8ZQK1_SALTY	Q90yf2	pleuronecte
135	31	91.2	179	2	Q90YF2_PLEPL	Q7lt82	bacterioph
136	31	91.2	208	2	Q7lt82_BPPI	Q500b3	pseudomonas
137	31	91.2	230	2	Q500B3_PSESY	Q4k3v9	pseudomonas
138	31	91.2	234	2	Q4K3V9_PSEFS	Q88vx4	lactobacill
139	31	91.2	235	2	Q88VX4_LACPL	Q5d9t2	schistosoma
140	31	91.2	241	2	Q5D9T2_SCHJA	Q87a46	xylella fas
141	31	91.2	257	1	Y1985_XYLFT	Q9paa9	xylella fas
142	31	91.2	257	1	Y2609_XYLFA	Q894d1	clostridium
143	31	91.2	265	2	Q894D1_CLOTE	Q9kbz5	bacillus ha
144	31	91.2	287	2	Q9KBZ5_BACHD	Q6l1b8	caenorhabdi
145	31	91.2	291	2	Q6L1B8_CAEBR	Q7m7w7	wolinella s
146	31	91.2	291	2	Q7M7W7_WOLSU	Q5mpq0	ovis aries
147	31	91.2	292	2	Q5MPQ0_SHEEP	Q5lt69	silicibacte
148	31	91.2	304	2	Q5LT69_SILPO	Q8nnz9	corynebacte
149	31	91.2	308	1	XERC_CORGL	Q8wv88	homo sapien
150	31	91.2	309	2	Q8WV88_HUMAN	Q5nyt6	azocarcus sp
151	31	91.2	311	2	Q5NYT6_AZOSE	Q978p0	thermoplasm
152	31	91.2	312	2	Q978P0_THEVO	Q4t3d9	tetraodon n
153	31	91.2	345	2	Q4T3D9_TETNG	Q5v5g5	haloarcula
154	31	91.2	359	2	Q5V5G5_HALMA	Q99kl7	mus musculus
155	31	91.2	362	2	Q99K17_MOUSE	Q7mrk7	wolinella s
156	31	91.2	365	2	Q7MRK7_WOLSU	Q88z21	lactobacill
157	31	91.2	368	2	Q88Z21_LACPL	Q07l77	rhodobacter
158	31	91.2	384	1	NIFS_RHOCA	Q4l1f8	burkholderi
159	31	91.2	391	2	Q4LLF8_9BURK	Q55km0	cryptococcu
160	31	91.2	406	2	Q55KM0_CRYNE	Q5kav1	cryptococcu
161	31	91.2	406	2	Q5KAV1_CRYNE	P15203	sus scrofa
162	31	91.2	409	1	TGFB3_PIG	P17125	mus musculus
163	31	91.2	410	1	TGFB3_MOUSE	Q7szv3	brachydanio
164	31	91.2	410	2	Q7SZV3_BRARE	Q66i23	brachydanio
165	31	91.2	410	2	Q66I23_BRARE	P16047	gallus gall
166	31	91.2	412	1	TGFB3_CHICK	P10600	homo sapien
167	31	91.2	412	1	TGFB3_HUMAN	Q07258	rattus norv
168	31	91.2	412	1	TGFB3_RAT	Q91yu7	mus musculus
169	31	91.2	412	2	Q91YU7_MOUSE	Q56a31	rattus norv
170	31	91.2	412	2	Q56A31_RAT	P00183	pseudomonas
171	31	91.2	414	1	CPXA_PSEPU	Q8x0i3	neurospora
172	31	91.2	426	2	Q8X0I3_NEUCR	Q582h7	trypanosoma
173	31	91.2	429	2	Q582H7_9TRYP	Q8dhl6	synechococc
174	31	91.2	440	2	Q8DHL6_SYNEL	Q4bxxv3	tetraodon n
175	31	91.2	442	2	Q4SXV3_TETNG	Q4rlv8	tetraodon n
176	31	91.2	444	2	Q4RLV8_TETNG	Q69bd7	campylobact
177	31	91.2	473	2	Q69BD7_CAMCO		

31	91.2	473	2	Q69B88_CAMJE	Q69b88	campylobact
31	91.2	486	2	Q63KP9_BURPS	Q63kp9	burkholderi
31	91.2	486	2	Q62CD2_BURMA	Q62cd2	burkholderi
31	91.2	488	2	Q93383_CABEL	Q93383	caenorhabdi
31	91.2	489	2	O62100_CABEL	O62100	caenorhabdi
31	91.2	554	2	Q92TT6_RHIME	Q92tt6	rhizobium m
31	91.2	563	2	Q60CM9_METCA	Q60cm9	methylococc
31	91.2	577	2	Q9QBQ5_9RETR	Q9qbq5	simian foam
31	91.2	600	2	Q7ZUZ1_BRARE	Q7zuz1	brachydanio
31	91.2	602	1	PEX5_HUMAN	P50542	homo sapien
31	91.2	602	2	Q91YC7_MOUSE	Q91yc7	mus musculu
31	91.2	603	2	Q920N3_CRIGR	Q920n3	cricetulus
31	91.2	608	2	Q642A7_RAT	Q642a7	rattus norv
31	91.2	615	2	Q4JUN3_CORJK	Q4jun3	corynebacte
31	91.2	631	2	Q96FN7_HUMAN	Q96fn7	homo sapien
31	91.2	633	2	Q55SW1_CRYNE	Q55aw1	cryptococcu
31	91.2	633	2	Q7TNJ1_CRIGR	Q7tnj1	cricetulus
31	91.2	635	2	Q5KH95_CRYNE	Q5kh95	cryptococcu
31	91.2	635	2	O30348_RALSO	O30348	ralstonia s
31	91.2	639	1	PEX5_MOUSE	O09012	mus musculu
31	91.2	639	2	Q8K2V5_MOUSE	Q8k2v5	mus musculu
31	91.2	640	2	O70525_CAVPO	O70525	cavia porce
31	91.2	640	2	Q920N5_CRIGR	Q920n5	cricetulus
31	91.2	645	2	Q5ZMQ9_CHICK	Q5zmq9	gallus gall
31	91.2	655	1	ACSA_NITEU	Q82ai5	nitrosomona
31	91.2	663	2	O81YJ8_HUMAN	Q81yj8	homo sapien
31	91.2	669	2	Q8DKV3_SYNEL	Q8dkv3	synechococc
31	91.2	677	2	Q4T3E0_TETNG	Q4t3e0	tetraodon n
31	91.2	717	2	Q6FVE9_CANGA	Q6fve9	candida gla
31	91.2	721	2	Q6BZZ6_YARLI	Q6bzz6	yarrowia li
31	91.2	738	2	Q4SNF4_TETNG	Q4snf4	tetraodon n
31	91.2	744	2	Q5WUJ4_LEGPL	Q5wuj4	legionella
31	91.2	744	2	Q5X334_LEGPA	Q5x334	legionella
31	91.2	744	2	Q5ZTB4_LEGPH	Q5ztb4	legionella
31	91.2	776	2	Q6K683_ORYSA	Q6k683	oryza sativ
31	91.2	848	2	Q4RVV3_TETNG	Q4rvv3	tetraodon n
31	91.2	863	2	Q9ST50_MAIZE	Q9st50	zea mays (m
31	91.2	865	2	Q8H0G9_TOBAC	Q8h0g9	nicotiana t
31	91.2	878	2	Q4SNF6_TETNG	Q4snf6	tetraodon n
31	91.2	936	2	Q9LPD9_ARATH	Q9lpd9	arabidopsis
31	91.2	944	2	Q5FQ83_GLUOX	Q5fq83	gluconobact
31	91.2	955	2	Q6QNH2_WHEAT	Q6qnh2	triticum ae
31	91.2	965	2	Q7Q3I3_ANOGA	Q7q3i3	anopheles g
31	91.2	968	2	Q6G5A6_BARHE	Q6g5a6	bartonella
31	91.2	980	2	Q83XJ4_VIBAN	Q83xj4	vibrio angu
31	91.2	984	2	Q9LZV1_ARATH	Q9lzv1	arabidopsis
31	91.2	1094	1	EMBA_MYCBO	P0a561	mycobacteri
31	91.2	1094	1	EMBA_MYCTU	P0a560	mycobacteri
31	91.2	1154	2	Q4QQD3_SCHMA	Q4qqd3	schistosoma
31	91.2	1230	2	Q889H2_PSESM	Q889h2	pseudomonas
31	91.2	1347	2	Q4LZ63_9BURK	Q4lzf3	burkholderi
31	91.2	1391	2	Q4N6K3_THEPA	Q4n6k3	theileria p
31	91.2	1413	2	Q4UJ10_THEAN	Q4uj10	theileria a
31	91.2	1447	2	Q4FWA4_LEIMA	Q4fwa4	leishmania
31	91.2	1600	2	Q4XWP5_PLACH	Q4xwp5	plasmodium
31	91.2	1633	2	Q51NH6_MAGGR	Q5inh6	magnaporthe
31	91.2	3006	2	Q9Y4D8_HUMAN	Q9y4d8	homo sapien
31	91.2	3680	2	Q5DIS7_PSEAE	Q5dis7	pseudomonas
31	91.2	3680	2	Q5DIU0_PSEAE	Q5diu0	pseudomonas
31	91.2	4342	2	Q9I157_PSEAE	Q9i157	pseudomonas
31	91.2	4889	2	Q4Z3Z9_PLABE	Q4z3z9	plasmodium
31	91.2	4971	2	Q8IBG1_PLAF7	Q8ibg1	plasmodium
31	91.2	5054	2	Q7RPJ1_PLAYO	Q7rpj1	plasmodium
30	88.2	45	2	Q91QU3_9GAMA	Q91qu3	human herpe
30	88.2	45	2	Q91QU5_9GAMA	Q91qu5	human herpe
30	88.2	46	2	Q91QS0_9GAMA	Q91qs0	pongine her
30	88.2	81	2	Q6U9H9_9CAUD	Q6u9h9	bacterioph
30	88.2	81	2	Q56EP0_9CAUD	Q56ep0	aeromonas p
30	88.2	91	2	Q7N8L3_PHOLL	Q7n8l3	photorhabdu
30	88.2	105	2	Q4LBJ6_PSESH	Q4lbj6	pseudomonas
30	88.2	105	2	Q4ZWF3_PSESY	Q4zwf3	pseudomonas
30	88.2	110	2	Q69141_9GAMA	Q69141	human herpe
30	88.2	120	1	E413_ADE12	P36709	human adeno

251	30	88.2	136	2	Q9DAL0_MOUSE	Q9dal0_mus musculus	324	30	88.2	346	2	Q8X8Z2_ECO57	Q8x8z2_escherichia
252	30	88.2	139	2	Q5TPZ4_ANOGA	Q5tpz4_anopheles g	325	30	88.2	350	2	Q6PJS6_HUMAN	Q6pjs6_homo sapien
253	30	88.2	156	2	Q7NWU0_CHRVO	Q7nwu0_chromobacte	326	30	88.2	353	2	Q6XN51_RHOER	Q6xn51_rhodococcus
254	30	88.2	158	2	Q9X0A5_THEMA	Q9x0a5_thermotoga	327	30	88.2	357	2	Q4R6X6_MACFA	Q4r6x6_macaca fasc
255	30	88.2	166	2	Q5DFY8_SCHJA	Q5dfy8_schistosoma	328	30	88.2	364	2	Q89W51_BRAJA	Q89w51_bradyrhizob
256	30	88.2	169	2	Q76LM7_HYDTH	Q76lm7_hydrogenoba	329	30	88.2	366	2	Q8XXM1_RALSO	Q8xxml_ralstonia s
257	30	88.2	176	2	Q5DA00_SCHJA	Q5da00_schistosoma	330	30	88.2	370	2	Q6N0D3_RHOPA	Q6n0d3_rhodopseudo
258	30	88.2	178	2	Q84XL1_LYCES	Q84xl1_lycopersico	331	30	88.2	372	2	Q8ZUR8_PYRAE	Q8zur8_pyrobaculum
259	30	88.2	181	2	Q6GCD0_STAAS	Q6gcd0_staphylococ	332	30	88.2	373	2	Q8C770_MOUSE	Q8c770_mus musculu
260	30	88.2	181	2	Q6GJX6_STAAR	Q6gjx6_staphylococ	333	30	88.2	373	2	Q9Z106_MOUSE	Q9z106_mus musculu
261	30	88.2	181	2	Q5HIV2_STAAC	Q5hiv2_staphylococ	334	30	88.2	376	2	Q6QIB7_9CAUD	Q6qib7_burkholderi
262	30	88.2	181	2	Q7A7M4_STAAN	Q7a7m4_staphylococ	335	30	88.2	376	2	Q4NFC4_MICC	Q4nfc4_arthrobacte
263	30	88.2	181	2	Q8NYA4_STAAM	Q8nya4_staphylococ	336	30	88.2	394	2	Q8UAR2_AGRF5	Q8uar2_agrobacteri
264	30	88.2	181	2	Q99WN5_STAAM	Q99wn5_staphylococ	337	30	88.2	401	2	Q4FXP7_LEIMA	Q4fxp7_leishmania
265	30	88.2	183	2	Q5XHD0_XENLA	Q5xhd0_xenopus lae	338	30	88.2	401	2	Q5X4W2_LEGPA	Q5x4w2_legionella
266	30	88.2	192	2	Q9VJY5_DROME	Q9vjy5_drosophila	339	30	88.2	406	1	IPSP_HUMAN	P05154_homo sapien
267	30	88.2	196	2	Q5RHX5_BRARE	Q5rhx5_brachydanio	340	30	88.2	406	2	Q4R6H4_MACFA	Q4r6h4_macaca fasc
268	30	88.2	197	2	Q7Z6B8_HUMAN	Q7z6b8_homo sapien	341	30	88.2	408	2	Q5YXZ5_NOCFA	Q5yxz5_nocardia fa
269	30	88.2	201	2	Q5NYK1_AZOSE	Q5nyk1_azocarcus sp	342	30	88.2	416	2	Q8WUF8_HUMAN	Q8wuf8_homo sapien
270	30	88.2	208	2	Q6PAM8_MOUSE	Q6pam8_mus musculu	343	30	88.2	416	2	Q5ZK44_CHICK	Q5zk44_gallus gall
271	30	88.2	210	2	Q6PTI9_9TURB	Q6pti9_stylochus s	344	30	88.2	420	2	Q82BK6_STRAW	Q82bk6_streptomyce
272	30	88.2	210	2	Q664A1_YERPS	Q664a1_yersinia ps	345	30	88.2	421	2	Q4SKG9_TETNG	Q4skg9_tetradodon n
273	30	88.2	224	2	Q4IF53_GIBZE	Q4if53_gibberella	346	30	88.2	425	2	Q6PFX1_MOUSE	Q6pfx1_mus musculu
274	30	88.2	229	2	Q45898_COXBU	Q45898_coxiella bu	347	30	88.2	427	2	Q9HU16_PSEAE	Q9hu16_pseudomonas
275	30	88.2	231	2	Q28500_MACMU	Q28500_macaca mula	348	30	88.2	433	2	Q9RD01_STRCO	Q9rd01_streptomyce
276	30	88.2	233	2	Q52877_COXBU	Q52877_coxiella bu	349	30	88.2	434	2	Q92T96_RHIME	Q92t96_rhizobium m
277	30	88.2	233	2	Q45912_COXBU	Q45912_coxiella bu	350	30	88.2	434	2	Q8DS32_STRMU	Q8ds32_streptococc
278	30	88.2	233	2	Q45935_COXBU	Q45935_coxiella bu	351	30	88.2	449	2	Q6FJM4_CANGA	Q6fjm4_candida gla
279	30	88.2	233	2	Q67SJ6_SYMTH	Q67sj6_symbiobacte	352	30	88.2	452	2	Q9YH86_CHICK	Q9yh86_gallus gall
280	30	88.2	233	2	Q83A05_COXBU	Q83a05_coxiella bu	353	30	88.2	466	2	Q53JJ2_ORYSA	Q53jj2_oryza sativ
281	30	88.2	236	2	Q5DK16_VIBAN	Q5dk16_vibrio angu	354	30	88.2	466	2	Q5FPU0_GLUOX	Q5fpu0_gluconobact
282	30	88.2	239	2	Q26768_METTH	Q26768_methanobact	355	30	88.2	469	2	Q8X1K0_COPCI	Q8x1k0_coprinus ci
283	30	88.2	252	2	Q8A4J0_BACTN	Q8a4j0_bacteroides	356	30	88.2	469	2	Q5P3L6_AZOSE	Q5p3l6_azocarcus sp
284	30	88.2	261	2	Q69147_9GAMA	Q69147_human herpe	357	30	88.2	472	2	Q6ATJ5_ORYSA	Q6atj5_oryza sativ
285	30	88.2	263	2	Q60407_CRICR	Q60407_cricetus cr	358	30	88.2	488	2	Q44317_9ARAC	Q44317_dugesieliella
286	30	88.2	272	2	Q829Z9_STRAW	Q829z9_streptomyce	359	30	88.2	501	2	Q8R9N8_THETN	Q8r9n8_thermoanaer
287	30	88.2	276	2	Q6UXP7_HUMAN	Q6uxp7_homo sapien	360	30	88.2	506	2	Q6FXL9_CANGA	Q6fxl9_candida gla
288	30	88.2	281	2	Q76D05_STRGR	Q76d05_streptomyce	361	30	88.2	506	2	Q69125_9GAMA	Q69125_human herpe
289	30	88.2	284	2	Q5CE87_CRYHO	Q5ce87_cryptospori	362	30	88.2	506	2	Q8AZK7_9GAMA	Q8azk7_human herpe
290	30	88.2	284	2	Q9HUX6_PSEAE	Q9hux6_pseudomonas	363	30	88.2	515	2	Q53470_9GCTU	Q53470_mycobacteri
291	30	88.2	289	2	Q8LB21_ARATH	Q8lb21_arabidopsis	364	30	88.2	515	2	Q6MAW6_PARUW	Q6maw6_parachlamyd
292	30	88.2	291	2	Q9GR92_ARTSF	Q9gr92_artemia san	365	30	88.2	520	2	Q5WAF7_BACSK	Q5waf7_bacillus cl
293	30	88.2	292	2	Q9SA78_ARATH	Q9sa78_arabidopsis	366	30	88.2	521	1	PGS1_SACPS	P79001_saccharomyc
294	30	88.2	292	2	Q8CCA0_MOUSE	Q8cca0_mus musculu	367	30	88.2	521	1	PGS1_YEAST	P25578_saccharomyc
295	30	88.2	301	2	Q4K8S1_PSEF5	Q4k8s1_pseudomonas	368	30	88.2	521	2	Q82507_ARATH	O82507_arabidopsis
296	30	88.2	301	2	Q9CWH0_MOUSE	Q9cwh0_mus musculu	369	30	88.2	521	2	Q82190_STRAW	Q82i90_streptomyce
297	30	88.2	305	2	Q5RHX6_BRARE	Q5rhx6_brachydanio	370	30	88.2	525	2	Q4P1M0_USTMA	Q4plm0_ustilago ma
298	30	88.2	306	2	Q4ZUQ4_PSESY	Q4zuq4_pseudomonas	371	30	88.2	527	2	Q5TVV8_ANOGA	Q5tlmv8_anopheles g
299	30	88.2	306	2	Q8EJU0_SHEON	Q8ej00_shewanella	372	30	88.2	527	2	Q8C9I4_MOUSE	Q8c9i4_mus musculu
300	30	88.2	306	2	Q883S3_PSESM	Q883s3_pseudomonas	373	30	88.2	531	2	Q65UK0_MANSM	Q65uk0_mannheimia
301	30	88.2	306	2	Q8C5X2_MOUSE	Q8c5x2_mus musculu	374	30	88.2	556	2	Q6XPA6_PHYIN	Q6xpa6_phytophthor
302	30	88.2	308	2	Q69136_9GAMA	Q69136_human herpe	375	30	88.2	556	2	Q6XPA7_PHYIN	Q6xpa7_phytophthor
303	30	88.2	309	2	Q92564_HUMAN	Q92564_homo sapien	376	30	88.2	556	2	Q6XPA9_9STRA	Q6xpa9_phytophthor
304	30	88.2	315	2	Q7Z3F3_HUMAN	Q7z3f3_homo sapien	377	30	88.2	556	2	Q6XPB0_9STRA	Q6xpb0_phytophthor
305	30	88.2	316	2	Q5FPU4_GLUOX	Q5fpu4_gluconobact	378	30	88.2	556	2	Q6XPB1_9STRA	Q6xpb1_phytophthor
306	30	88.2	318	2	Q5BEU1_EMENI	Q5beu1_aspergillus	379	30	88.2	556	2	Q6XPB2_9STRA	Q6xpb2_phytophthor
307	30	88.2	318	2	Q5H033_XANOR	Q5h033_xanthomonas	380	30	88.2	556	2	Q6XPB3_9STRA	Q6xpb3_phytophthor
308	30	88.2	318	2	Q87D99_XYLFT	Q87d99_xylella fas	381	30	88.2	556	2	Q6XPB4_PHYIN	Q6xpb4_phytophthor
309	30	88.2	318	2	Q8PM38_XANAC	Q8pm38_xanthomonas	382	30	88.2	556	2	Q6XPB5_PHYIN	Q6xpb5_phytophthor
310	30	88.2	319	2	Q9PBQ1_XYLFA	Q9pbq1_xylella fas	383	30	88.2	556	2	Q6XPB6_PHYIN	Q6xpb6_phytophthor
311	30	88.2	321	2	Q4UT87_XANCP	Q4ut87_xanthomonas	384	30	88.2	556	2	Q6XPB7_PHYIN	Q6xpb7_phytophthor
312	30	88.2	321	2	Q8PAD8_XANCP	Q8pad8_xanthomonas	385	30	88.2	556	2	Q6XPB8_PHYIN	Q6xpb8_phytophthor
313	30	88.2	323	1	LRP16_MOUSE	Q922b1_mus musculu	386	30	88.2	556	2	Q6XPB9_PHYIN	Q6xpb9_phytophthor
314	30	88.2	326	2	Q7WXM1_ALCEU	Q7wxm1_alcaligenes	387	30	88.2	556	2	Q6XPC0_PHYIN	Q6xpc0_phytophthor
315	30	88.2	326	2	Q4LRI4_9BURK	Q4lri4_burkholderi	388	30	88.2	556	2	Q6XPC1_PHYIN	Q6xpc1_phytophthor
316	30	88.2	327	2	Q5LR05_SILPO	Q5lro5_silicibacte	389	30	88.2	556	2	Q6XPC3_9STRA	Q6xpc3_phytophthor
317	30	88.2	329	2	Q7SB66_NEUCR	Q7sb66_neurospora	390	30	88.2	556	2	Q6XPC5_PHYIN	Q6xpc5_phytophthor
318	30	88.2	332	1	EXPA_CEPAC	P11935_c_cephalospor	391	30	88.2	556	2	Q6XPC6_PHYIN	Q6xpc6_phytophthor
319	30	88.2	332	2	Q9P4T5_CEPAC	Q9p4t5_cephalospor	392	30	88.2	556	2	Q84NE6_PHYIN	Q84ne6_phytophthor
320	30	88.2	334	2	Q5V1K1_HALMA	Q5v1k1_haloarcula	393	30	88.2	561	2	Q4PGB7_USTMA	Q4pgb7_ustilago ma
321	30	88.2	338	2	Q9EYD4_ECO57	Q9eyd4_escherichia	394	30	88.2	568	2	Q68HC9_SOLTU	Q68hc9_solanum tub
322	30	88.2	338	2	Q4LPN5_9BURK	Q4lpn5_burkholderi	395	30	88.2	569	2	Q6P3G2_BRARE	Q6p3g2_brachydanio
323	30	88.2	342	2	Q5YMK5_NOCFA	Q5ymk5_nocardia fa	396	30	88.2	577	2	Q86DC0_CAEEL	Q86dc0_caenorhabdi

397	30	88.2	595	2	Q716S3_9CAUD	Q716s3 mycobacteri	470	30	88.2	1900	2	Q8WTD2_DROME	Q8wtd2 drosophila
398	30	88.2	601	2	Q5CTA2_CRYPV	Q5cta2 cryptospori	471	30	88.2	1900	2	Q9VTS9_DROME	Q9vts9 drosophila
399	30	88.2	601	2	Q6D0G9_ERWCT	Q6d0g9 erwinia car	472	30	88.2	1915	2	Q5VXF6_HUMAN	Q5vxf6 homo sapien
400	30	88.2	602	2	Q6CF71_YARLI	Q6cf71 yarrowia li	473	30	88.2	1947	2	Q4SIP4_TETNG	Q4sip4 tetraodon n
401	30	88.2	608	2	Q8QZW3_MOUSE	Q8qzw3 mus musculu	474	30	88.2	2072	2	Q4SAO9_TETNG	Q4sa09 tetraodon n
402	30	88.2	613	1	GLSA1_BRAJA	Q89na7 bradyrhizob	475	30	88.2	2145	2	Q4RHK4_TETNG	Q4rkh4 tetraodon n
403	30	88.2	614	2	Q69HN8_CIOIN	Q69hn8 ciona intes	476	30	88.2	2179	2	Q91DM0_9VIRU	Q91dm0 petunia vei
404	30	88.2	614	2	Q6D6M5_ERWCT	Q6d6m5 erwinia car	477	30	88.2	2180	2	Q6XKE6_9VIRU	Q6xke6 petunia vei
405	30	88.2	619	2	Q73SI1_MYCPA	Q73si1 mycobacteri	478	30	88.2	2202	2	Q59EQ7_HUMAN	Q59eq7 homo sapien
406	30	88.2	628	2	Q5JM56_ORYSA	Q5jm56 oryza sativ	479	30	88.2	2592	2	Q84T10_ORYSA	Q84t10 oryza sativ
407	30	88.2	632	2	Q75IJ9_ORYSA	Q75ij9 oryza sativ	480	30	88.2	3038	1	TRIO_HUMAN	Q75962 homo sapien
408	30	88.2	632	2	Q57K30_SALCH	Q57k30 salmonella	481	30	88.2	3072	2	Q6LEZ2_PLAF7	Q6lez2 plasmodium
409	30	88.2	639	2	Q9Z4I7_STRSQ	Q9z4i7 streptomyce	482	30	88.2	3470	2	Q88F79_PSEPK	Q88f79 pseudomonas
410	30	88.2	639	2	Q8K0M2_MOUSE	Q8k0m2 mus musculu	483	30	88.2	4034	2	Q7R5J5_GIALA	Q7r5j5 giardia lam
411	30	88.2	656	2	Q4Q0K8_LEIMA	Q4q0k8 leishmania	484	30	88.2	6260	2	Q54299_STRHY	Q54299 streptomyce
412	30	88.2	658	1	SPEA_ECO57	Q8xcx9 escherichia	485	30	88.2	8348	2	Q4S6Y5_TETNG	Q4s6y5 tetraodon n
413	30	88.2	658	1	SPEA_ECOL6	Q8fe34 escherichia	486	30	88.2	8563	2	Q54297_STRHY	Q54297 streptomyce
414	30	88.2	658	1	SPEA_ECOLI	P21l70 escherichia	487	29	85.3	42	2	Q8EJZ9_SHEON	Q8ejz9 shewanella
415	30	88.2	658	1	SPEA_SALTI	P60658 salmonella	488	29	85.3	67	2	Q51A41_ENTHI	Q51a41 entamoeba h
416	30	88.2	658	1	SPEA_SALTY	P60659 salmonella	489	29	85.3	74	2	Q7VK46_HELHP	Q7vk46 helicobacte
417	30	88.2	659	1	SPEA_YERPE	Q8zhg8 yersinia pe	490	29	85.3	80	2	Q9KRZ6_VIBCH	Q9krz6 vibrio chol
418	30	88.2	661	2	Q6AX23_XENLA	Q6ax23 xenopus lae	491	29	85.3	81	2	Q6EB44_CAMJE	Q6eb44 campylobact
419	30	88.2	662	1	SPEA_SHIFL	Q83q93 shigella fl	492	29	85.3	84	2	O80132_9CAUD	O80132 bacterioph
420	30	88.2	662	2	Q4WU83_ASPPU	Q4wu83 aspergillus	493	29	85.3	85	2	O21893_9CAUD	O21893 bacterioph
421	30	88.2	682	2	Q8BK93_MOUSE	Q8bk93 mus musculu	494	29	85.3	85	2	Q8FN33_COREF	Q8fn33 corynebacte
422	30	88.2	706	2	Q8VNN5_9VIRU	Q8vnn5 bacterioph	495	29	85.3	87	2	Q9DFJ7_GILMI	Q9dfj7 gillichthys
423	30	88.2	712	2	Q6C6Z8_YARLI	Q6c6z8 yarrowia li	496	29	85.3	90	1	Y4IG_RHISN	P55490 rhizobium s
424	30	88.2	713	2	Q18662_CAEEL	Q18662 caenorhabdi	497	29	85.3	94	2	Q63NA3_BURPS	Q63na3 burkholderi
425	30	88.2	733	2	Q5XQF4_MAIZE	Q5xqf4 zea mays (m	498	29	85.3	95	2	Q86SR9_HUMAN	Q86sr9 homo sapien
426	30	88.2	736	2	Q8KNK0_SALTI	Q8knk0 salmonella	499	29	85.3	99	2	Q7NJ12_GLOVI	Q7nj12 gloeobacter
427	30	88.2	745	2	Q4S3N3_TETNG	Q4s3n3 tetraodon n	500	29	85.3	99	2	Q7NMC1_GLOVI	Q7nmc1 gloeobacter
428	30	88.2	751	2	Q6BXA9_DEBHA	Q6bxa9 debaryomyce	501	29	85.3	102	2	Q4KA72_PSEF5	Q4ka72 pseudomonas
429	30	88.2	756	2	Q4IZG3_AZОВI	Q4izg3 azotobacter	502	29	85.3	102	2	Q9J9C0_9LUTE	Q9j9c0 sugarcane y
430	30	88.2	774	2	Q7QCC9_ANOGA	Q7qcc9 anopheles g	503	29	85.3	106	1	Y04K_BFT4	P07069 bacterioph
431	30	88.2	781	2	Q5AZZ3_EMENI	Q5azz3 aspergillus	504	29	85.3	109	2	Q8KUW1_SYNP7	Q8kuw1 synechococc
432	30	88.2	830	2	Q5YRU2_NOCFA	Q5yru2 nocardia fa	505	29	85.3	114	2	Q7BF85_YERPS	Q7bf85 yersinia ps
433	30	88.2	897	2	Q7U1L0_MYCBO	Q7u1l0 mycobacteri	506	29	85.3	114	2	Q93PD5_YERPE	Q93pd5 yersinia pe
434	30	88.2	912	2	Q8CCJ2_MOUSE	Q8ccj2 mus musculu	507	29	85.3	114	2	Q66A07_YERPS	Q66a07 yersinia ps
435	30	88.2	923	2	Q5E1V0_VIBF1	Q5e1v0 vibrio fisc	508	29	85.3	115	2	Q8S0G0_ORYSA	Q8s0g0 oryza sativ
436	30	88.2	952	2	Q9LQ93_ARATH	Q9lq93 arabidopsis	509	29	85.3	116	2	Q7P7E0_FUSNV	Q7p7e0 fusobacteri
437	30	88.2	975	2	Q4WNK0_ASPPU	Q4wnk0 aspergillus	510	29	85.3	120	2	Q8CL76_YERPE	Q8cl76 yersinia pe
438	30	88.2	1009	2	Q7SEY7_NEUCR	Q7sey7 neurospora	511	29	85.3	121	2	Q695Q6_9ADEN	Q695q6 simian aden
439	30	88.2	1026	2	Q86T76_HUMAN	Q86t76 homo sapien	512	29	85.3	121	2	Q6QP86_9ADEN	Q6qp86 simian aden
440	30	88.2	1027	1	DOCK1_MOUSE	Q8bur4 mus musculu	513	29	85.3	121	2	Q6QPC2_9ADEN	Q6qpc2 simian aden
441	30	88.2	1033	2	Q4J2V5_AZОВI	Q4j2v5 azotobacter	514	29	85.3	121	2	Q6QPF8_9ADEN	Q6qpf8 simian aden
442	30	88.2	1040	2	Q9T014_ARATH	Q9t014 arabidopsis	515	29	85.3	121	2	Q8UY65_9ADEN	Q8uy65 simian aden
443	30	88.2	1083	2	Q9RPH6_MYCSM	Q9rph6 mycobacteri	516	29	85.3	122	2	Q6H1A5_9ADEN	Q6h1a5 human adeno
444	30	88.2	1094	1	EX5B_MYCTU	P96920 mycobacteri	517	29	85.3	122	2	Q5VH99_ADE04	Q5vh99 human adeno
445	30	88.2	1099	2	Q571B9_MOUSE	Q571b9 mus musculu	518	29	85.3	124	2	Q8U6I8_AGR75	Q8u6i8 agrobacteri
446	30	88.2	1108	2	Q4SEA7_TETNG	Q4sea7 tetraodon n	519	29	85.3	126	2	Q7QYD0_GIALA	Q7qyd0 giardia lam
447	30	88.2	1123	2	Q43948_PLAFA	Q43948 plasmodium	520	29	85.3	131	2	Q937P9_SALET	Q937p9 salmonella
448	30	88.2	1189	2	Q7URB6_RHOBA	Q7urb6 rhodopirell	521	29	85.3	131	2	Q70DP1_ACTPL	Q70dp1 actinobacil
449	30	88.2	1206	2	Q4IM26_GIBZE	Q4im26 gibberella	522	29	85.3	131	2	Q70DP2_ACTLI	Q70dp2 actinobacil
450	30	88.2	1294	2	Q4RFA4_TETNG	Q4rfa4 tetraodon n	523	29	85.3	133	2	Q4ZPV7_PSESY	Q4zpv7 pseudomonas
451	30	88.2	1334	2	Q4RTS6_TETNG	Q4rts6 tetraodon n	524	29	85.3	134	2	Q886I5_PSESM	Q886i5 pseudomonas
452	30	88.2	1364	2	Q8PJN8_XANAC	Q8pjn8 xanthomonas	525	29	85.3	137	2	Q50LS9_ENTHI	Q50ls9 entamoeba h
453	30	88.2	1402	2	Q5GZ20_XANOR	Q5gz20 xanthomonas	526	29	85.3	143	2	O59287_PYRHO	O59287 pyrococcus
454	30	88.2	1424	2	Q6P9K6_MOUSE	Q6p9k6 mus musculu	527	29	85.3	145	2	Q5VPU6_ORYSA	Q5vpu6 oryza sativ
455	30	88.2	1566	2	Q8MZD0_DROME	Q8mzd0 drosophila	528	29	85.3	145	2	Q5P0R6_AZOSE	Q5p0r6 azoarcus sp
456	30	88.2	1566	2	Q9VJT4_DROME	Q9vjt4 drosophila	529	29	85.3	149	2	Q9Z2R2_MOUSE	Q9z2r2 mus musculu
457	30	88.2	1569	2	Q8WTD1_DROME	Q8wtd1 drosophila	530	29	85.3	155	2	Q4WLQ8_ASPPU	Q4wlq8 aspergillus
458	30	88.2	1569	2	Q9VTS8_DROME	Q9vts8 drosophila	531	29	85.3	155	2	Q834C9_ENTFA	Q834c9 enterococcu
459	30	88.2	1577	2	Q9NKC7_DROME	Q9nkc7 drosophila	532	29	85.3	156	2	Q8LDX8_ARATH	Q8ldx8 arabidopsis
460	30	88.2	1606	2	Q9XB15_MYCBO	Q9xb15 mycobacteri	533	29	85.3	158	2	Q6U9T5_9CAUD	Q6u9t5 bacterioph
461	30	88.2	1606	2	Q7D7L9_MYCTU	Q7d7l9 mycobacteri	534	29	85.3	158	2	Q56EZ5_9CAUD	Q56ez5 aeromonas p
462	30	88.2	1606	2	Q7TZ47_MYCBO	Q7tz47 mycobacteri	535	29	85.3	159	2	Q8SS75_ENCCU	Q8ss75 encephalito
463	30	88.2	1653	2	Q5YM84_NOCFA	Q5ym84 nocardia fa	536	29	85.3	159	2	Q67TG9_SYMTH	Q67tg9 symbiobacte
464	30	88.2	1669	2	Q7N7D7_PHOLL	Q7n7d7 photorhabdu	537	29	85.3	161	2	O5B6X4_EMENI	O5b6x4 aspergillus
465	30	88.2	1670	2	Q8WS23_DROME	Q8ws23 drosophila	538	29	85.3	162	2	Q7N9X2_PHOLL	Q7n9x2 photorhabdu
466	30	88.2	1670	2	Q9VTS6_DROME	Q9vts6 drosophila	539	29	85.3	163	2	Q873A6_NEUCR	Q873a6 neurospora
467	30	88.2	1865	1	DOCK1_HUMAN	Q14i85 homo sapien	540	29	85.3	165	2	O66637_AQUAE	O66637 aquifex aeo
468	30	88.2	1886	2	Q6XKE5_9VIRU	Q6xke5 petunia vei	541	29	85.3	169	2	Q4I938_GIBZE	Q4i938 gibberella
469	30	88.2	1900	2	Q8WS24_DROME	Q8ws24 drosophila	542	29	85.3	170	2	Q51MR3_MAGGR	Q51mr3 magnaporthe

543	29	85.3	170	2	Q7N6G7_PHOLL	Q7n6g7_photorhabdu	616	29	85.3	244	2	Q7WP46_BORBR	Q7wp46_bordetella
544	29	85.3	170	2	Q7NS63_CHRVO	Q7ns63_chromobacte	617	29	85.3	244	2	Q4V8S1_BRARE	Q4v8s1_brachydanio
545	29	85.3	172	2	Q5E537_VIBF1	Q5e537_vibrio fisc	618	29	85.3	245	2	Q8EPY9_OCEIH	Q8epy9_oceanobacil
546	29	85.3	172	2	Q63KS2_BURPS	Q63ks2_burkholderi	619	29	85.3	246	2	Q4IVT5_AZOVI	Q4ivt5_azotobacter
547	29	85.3	174	2	Q7ND96_GLOVI	Q7nd96_gloeobacter	620	29	85.3	248	2	Q4TKN9_9SPHN	Q4tkn9_erythrobact
548	29	85.3	174	2	Q64683_MESAU	Q64683_mesocricetu	621	29	85.3	248	2	Q63VF1_BURPS	Q63vf1_burkholderi
549	29	85.3	175	2	Q4L552_STAHI	Q4l552_staphylococ	622	29	85.3	248	2	Q90XF9_BRARE	Q90xf9_brachydanio
550	29	85.3	175	2	Q7MV46_PORGI	Q7mv46_porphyromon	623	29	85.3	250	2	Q746K8_THET2	Q746k8_thermus the
551	29	85.3	182	2	Q8EKZ1_OCEIH	Q8ekz1_oceanobacil	624	29	85.3	250	2	Q53W72_THET8	Q53w72_thermus the
552	29	85.3	184	2	Q9ZIT6_ECOLI	Q9zit6_escherichia	625	29	85.3	251	2	Q4I3H5_GIBZE	Q4i3h5_gibberella
553	29	85.3	186	2	Q4NNA7_9DELT	Q4nna7_anaeromyxob	626	29	85.3	251	2	Q9X7N9_STRCO	Q9x7n9_streptomyce
554	29	85.3	189	2	Q6IMH5_NEOCA	Q6imh5_neospora ca	627	29	85.3	252	2	Q82PM0_STRAW	Q82pm0_streptomyce
555	29	85.3	192	1	MOBA_NEIMA	Q9juas5_neisseria m	628	29	85.3	252	2	Q8FNW6_COREF	Q8fnw6_corynebacte
556	29	85.3	192	1	MOBA_NEIMB	P58747_neisseria m	629	29	85.3	254	2	Q77K62_9NUCL	Q77k62_helicoverpa
557	29	85.3	192	2	Q7QW59_GIALA	Q7qw59_giardia lam	630	29	85.3	254	2	Q77LX4_9NUCL	Q77lx4_helicoverpa
558	29	85.3	192	2	Q7NT62_CHRVO	Q7nt62_chromobacte	631	29	85.3	254	2	Q99GW8_9NUCL	Q99gw8_helicoverpa
559	29	85.3	193	2	Q7NV13_CHRVO	Q7nvi3_chromobacte	632	29	85.3	255	1	Y1632_METJA	Q59026_methanococc
560	29	85.3	196	2	Q5P7P9_AZOSE	Q5p7p9_azoarcus sp	633	29	85.3	258	2	Q88PQ7_PSEPK	Q88pq7_pseudomonas
561	29	85.3	197	2	Q8PN82_XANAC	Q8pn82_xanthomonas	634	29	85.3	263	2	Q4KEM0_PSEF5	Q4kem0_pseudomonas
562	29	85.3	198	2	Q8XZR6_RALSO	Q8xizr6_ralstonia s	635	29	85.3	263	2	Q88J12_PSEPK	Q88ji2_pseudomonas
563	29	85.3	199	1	TORD_ECOL6	Q8fj52_escherichia	636	29	85.3	266	2	Q5XGF8_XENTR	Q5xgf8_xenopus tro
564	29	85.3	199	2	Q5N2X9_SYNP6	Q5n2x9_synechococc	637	29	85.3	266	2	Q6NU97_XENLA	Q6nu97_xenopus lae
565	29	85.3	199	2	Q42037_9GAMA	Q42037_alcelaphine	638	29	85.3	269	1	RIBF_MYCPN	P75587_mycoplasma
566	29	85.3	201	2	Q7CMH7_BACAN	Q7cmh7_bacillus an	639	29	85.3	269	2	Q5XHE1_XENLA	Q5xhe1_xenopus lae
567	29	85.3	201	2	Q4MKL0_BACCE	Q4mk10_bacillus ce	640	29	85.3	270	2	Q4TAY6_TETNG	Q4tay6_tetraodon n
568	29	85.3	201	2	Q74P26_BACC1	Q74p26_bacillus ce	641	29	85.3	271	2	Q4W477_AERSA	Q4w477_aeromonas s
569	29	85.3	201	2	Q82VG6_NITEU	Q82vg6_nitrosomona	642	29	85.3	271	2	Q699Q3_AERHY	Q699q3_aeromonas h
570	29	85.3	201	2	Q9X317_BACAN	Q9x317_bacillus an	643	29	85.3	271	2	Q5YY59_NOCFA	Q5yy59_nocardia fa
571	29	85.3	202	2	Q75AD3_ASHGO	Q75ad3_ashbya goss	644	29	85.3	273	2	Q67LW2_SYMTH	Q67lw2_symbiobacte
572	29	85.3	204	2	Q5F818_NEIG1	Q5f818_neisseria g	645	29	85.3	274	2	Q6P0C0_BRARE	Q6p0c0_brachydanio
573	29	85.3	204	2	Q67KU0_SYMTH	Q67ku0_symbiobacte	646	29	85.3	275	2	Q6NEN3_CORDI	Q6nen3_corynebacte
574	29	85.3	206	2	Q4T545_TETNG	Q4t545_tetraodon n	647	29	85.3	276	2	Q4YW93_PLABE	Q4yw93_plasmodium
575	29	85.3	208	2	Q67TGO_SYMTH	Q67tgo_symbiobacte	648	29	85.3	277	2	Q58047_PYROCOC	O58047_pyrococcus
576	29	85.3	210	2	Q8VP38_ECOLI	Q8vp38_escherichia	649	29	85.3	279	2	Q86LQ0_BRABE	Q86lq0_branchiosto
577	29	85.3	210	2	Q76BF5_ORYLA	Q76bf5_oryzias lat	650	29	85.3	280	2	Q5N664_SYNP6	Q5n664_synechococc
578	29	85.3	211	2	Q70I46_DIPPU	Q70i46_diptoptera	651	29	85.3	283	1	Y1654_HAEIN	P45298_haemophilus
579	29	85.3	211	2	Q7NL44_GLOVI	Q7nl44_gloeobacter	652	29	85.3	283	2	O85856_SPHAR	O85856_sphingomona
580	29	85.3	213	2	Q5DDT7_SCHJA	Q5ddt7_schistosoma	653	29	85.3	283	2	Q4QJT9_HAEI8	Q4qjt9_haemophilus
581	29	85.3	214	1	Y220B_HAEIN	O86222_haemophilus	654	29	85.3	284	2	Q4KPH2_PSEF5	Q4kfh2_pseudomonas
582	29	85.3	214	2	Q4QNW7_CHRVO	Q4qnw7_haemophilus	655	29	85.3	285	2	Q9HLI1_THEAC	Q9hli1_thermoplasm
583	29	85.3	214	2	Q7NVB7_CHRVO	Q7nvb7_chromobacte	656	29	85.3	285	2	Q4J2D1_AZOVI	Q4j2d1_azotobacter
584	29	85.3	215	2	Q8NTE3_CORGL	Q8nte3_corynebacte	657	29	85.3	285	2	Q9Z867_CHLPN	Q9z867_chlamydia p
585	29	85.3	216	1	PAAD_STRCO	Q9kyp1_streptomyce	658	29	85.3	286	2	Q89V19_BRAJA	Q89v19_bradyrhizob
586	29	85.3	216	2	Q4WUZ8_ASPFU	Q4wuz8_aspergillus	659	29	85.3	287	2	Q4ZVF7_PSESY	Q4zvf7_pseudomonas
587	29	85.3	220	2	Q4R883_MACFA	Q4r883_macaca fasc	660	29	85.3	288	1	ZIPA_PSESM	Q87YY5_pseudomonas
588	29	85.3	220	2	Q5YUZ0_NOCFA	Q5yuz0_nocardia fa	661	29	85.3	289	1	ZIPA_PSEAE	Q9i3i5_pseudomonas
589	29	85.3	221	2	Q67M46_SYMTH	Q67m46_symbiobacte	662	29	85.3	289	2	Q4S363_TETNG	Q4s363_tetraodon n
590	29	85.3	223	2	Q7VSF0_BORPE	Q7vsf0_bordetella	663	29	85.3	290	2	O5BJN1_RAT	Q5bjn1_rattus norv
591	29	85.3	223	2	Q7W5H5_BORPA	Q7w5h5_bordetella	664	29	85.3	291	1	PCTL_HUMAN	Q9Y365_homo sapien
592	29	85.3	223	2	Q7WD11_BORBR	Q7wd11_bordetella	665	29	85.3	291	1	PCTL_MOUSE	Q9jmd3_mus musculu
593	29	85.3	224	2	Q88LM3_PSEPK	Q88lm3_pseudomonas	666	29	85.3	295	2	Q4TNL4_9SPHN	Q4tnl4_erythrobact
594	29	85.3	226	2	Q62CA9_BURNA	Q62ca9_burkholderi	667	29	85.3	296	1	KSGA_STAEP	Q8cqu5_s dimethyla
595	29	85.3	229	2	Q7P6L7_FUSNV	Q7p6l7_fusobacteri	668	29	85.3	296	1	KSGA_STAEO	Q5hrr2_s dimethyla
596	29	85.3	229	2	Q4HBD4_9DEIO	Q4hbd4_deinococcus	669	29	85.3	297	1	ZIPA_PSEPK	Q4n4c6_theileria p
597	29	85.3	229	2	Q8REW9_FUSNN	Q8rew9_fusobacteri	670	29	85.3	297	1	Q5YPT6_NOCFA	Q88f24_pseudomonas
598	29	85.3	231	2	Q5GY89_XANOR	Q5gy89_xanthomonas	671	29	85.3	298	2	Q89E74_BRAJA	Q5ypt6_nocardia fa
599	29	85.3	231	2	Q82MI2_STRAW	Q82mi2_streptomyce	672	29	85.3	299	2	Q9L526_VIBCH	Q89e74_bradyrhizob
600	29	85.3	232	2	Q5YQZ4_NOCFA	Q5yqz4_nocardia fa	673	29	85.3	301	2	Q4NHT3_9MICC	Q9l526_vibrio chol
601	29	85.3	232	2	Q9HTS3_PSEAE	Q9hts3_pseudomonas	674	29	85.3	301	2	Q9KKY6_VIBCH	Q4nht3_arthrobacte
602	29	85.3	233	2	Q8PQP0_XANAC	Q8ppq0_xanthomonas	675	29	85.3	301	2	Q86XV4_HUMAN	Q9kky6_vibrio chol
603	29	85.3	236	2	Q57YPI_9TRYP	Q57yp1_trypanosoma	676	29	85.3	305	2	Q7W2V9_BORPA	Q86xv4_homo sapien
604	29	85.3	238	2	Q835J7_ENTFA	Q835j7_enterococcu	677	29	85.3	307	2	Q7WDV9_BORBR	Q7w2v9_bordetella
605	29	85.3	238	2	Q67QU0_SYMTH	Q67qu0_symbiobacte	678	29	85.3	307	2	Q7MIJ5_PYLLI	Q7wdv9_bordetella
606	29	85.3	239	2	Q92387_TRIVR	Q92387_trigonopsis	679	29	85.3	308	2	Q4NG45_9MICC	Q7mlj5_pylaiella l
607	29	85.3	239	2	Q4LS99_9BURK	Q4ls99_burkholderi	680	29	85.3	308	2	Q4NG45_9MICC	Q4ng45_arthrobacte
608	29	85.3	239	2	Q7W1D9_BORPA	Q7w1d9_bordetella	681	29	85.3	308	2	Q6LG31_PHOPR	Q6lg31_photobacter
609	29	85.3	241	2	Q7P147_CHRVO	Q7p147_chromobacte	682	29	85.3	309	2	Q7Y510_9CAUD	Q7y510_xanthomonas
610	29	85.3	242	2	Q8TXS6_METKA	Q8txs6_methanopyru	683	29	85.3	309	2	Q7WBS3_BORPA	Q7wbs3_bordetella
611	29	85.3	243	2	Q45899_COXBU	Q45899_coxiella bu	684	29	85.3	309	2	Q7WPR6_BORBR	Q7wpr6_bordetella
612	29	85.3	243	2	Q45925_COXBU	Q45925_coxiella bu	685	29	85.3	311	1	GPDA_HELHP	Q7vil5_helicobacte
613	29	85.3	243	2	Q9S618_COXBU	Q9s618_coxiella bu	686	29	85.3	313	2	Q8XRE2_RALSO	Q8xre2_ralstonia s
614	29	85.3	243	2	Q45943_COXBU	Q45943_coxiella bu	687	29	85.3	315	2	Q4HVD0_GIBZE	Q4hvd0_gibberella
615	29	85.3	244	2	Q7W0F8_BORPE	Q7w0f8_bordetella	688	29	85.3	318	2	Q4GYV8_9TRYP	Q4gyv8_trypanosoma

616	29	85.3	244	2	Q7WP46_BORBR	Q7wp46_bordetella
617	29	85.3	244	2	Q4V8S1_BRARE	Q4v8s1_brachydanio
618	29	85.3	245	2	Q8EPY9_OCEIH	Q8epy9_oceanobacil
619	29	85.3	246	2	Q4IVT5_AZOVI	Q4ivt5_azotobacter
620	29	85.3	248	2	Q4TKN9_9SPHN	Q4tkn9_erythrobact
621	29	85.3	248	2	Q63VF1_BURPS	Q63vf1_burkholderi
622	29	85.3	248	2	Q90XF9_BRARE	Q90xf9_brachydanio
623	29	85.3	250	2	Q746K8_THET2	Q746k8_thermus the
624	29	85.3	250	2	Q53W72_THET8	Q53w72_thermus the
625	29	85.3	251	2	Q4I3H5_GIBZE	Q4i3h5_gibberella
626	29	85.3	251	2	Q9X7N9_STRCO	Q9x7n9_streptomyce
627	29	85.3	252	2	Q82PM0_STRAW	Q82pm0_streptomyce
628	29	85.3	252	2	Q8FNW6_COREF	Q8fnw6_corynebacte
629	29	85.3	254	2	Q77K62_9NUCL	Q77k62_helicoverpa
630	29	85.3	254	2	Q77LX4_9NUCL	Q77lx4_helicoverpa
631	29	85.3	254	2	Q99GW8_9NUCL	Q99gw8_helicoverpa
632	29	85.3	255	1	Y1632_METJA	Q59026_methanococc
633	29	85.3	258	2	Q88PQ7_PSEPK	Q88pq7_pseudomonas
634	29	85.3	263	2	Q4KEM0_PSEF5	Q4kem0_pseudomonas
635	29	85.3	263	2	Q88J12_PSEPK	Q88ji2_pseudomonas
636	29	85.3	266	2	Q5XGF8_XENTR	Q5xgf8_xenopus tro
637	29	85.3	266	2	Q6NU97_XENLA	Q6nu97_xenopus lae
638	29	85.3	269	1	RIBF_MYCPN	P75587_mycoplasma
639	29	85.3	269	2	Q5XHE1_XENLA	Q5xhe1_xenopus lae
640	29	85.3	270	2	Q4TAY6_TETNG	Q4tay6_tetraodon n
641	29	85.3	271	2	Q4W477_AERSA	Q4w477_aeromonas s
642	29	85.3	271	2	Q699Q3_AERHY	Q699q3_aeromonas h
643	29	85.3	271	2	Q5YY59_NOCFA	Q5yy59_nocardia fa
644	29	85.3	273	2	Q67LW2_SYMTH	Q67lw2_symbiobacte
645	29	85.3	274	2	Q6P0C0_BRARE	Q6p0c0_brachydanio
646	29	85.3	275	2	Q6NEN3_CORDI	Q6nen3_corynebacte
647	29	85.3	276	2	Q4YW93_PLABE	Q4yw93_plasmodium
648	29	85.3	277	2	O58047_PYRHO	O58047_pyrococcus
649	29	85.3	279	2	Q86LQ0_BRABE	Q86lq0_branchiosteo
650	29	85.3	280	2	Q5N664_SYNP6	Q5n664_synechococc
651	29	85.3	283	1	Y1654_HAEIN	P45298_haemophilus
652	29	85.3	283	2	Q85856_SPHAR	Q85856_sphingomonas
653	29	85.3	283	2	Q4QJT9_HAEI8	Q4qjt9_haemophilus
654	29	85.3	284	2	Q4KPH2_PSEF5	Q4kfh2_pseudomonas
655	29	85.3	285	2	Q9HLI1_THEAC	Q9hli1_thermoplasma
656	29	85.3	285	2	Q4J2D1_AZOVI	Q4j2d1_azotobacter
657	29	85.3	285	2	Q9Z867_CHLPN	Q9z867_chlamydia p
658	29	85.3	286	2	Q89V19_BRAJA	Q89v19_bradyrhizob
659	29	85.3	287	2	Q4ZVF7_PSESY	Q4zvf7_pseudomonas
660	29	85.3	288	1	ZIPA_PSESM	Q87yy5_pseudomonas
661	29	85.3	289	1	ZIPA_PSEAE	Q9i3i5_pseudomonas
662	29	85.3	289	2	Q4S363_TETNG	Q9i3i5_pseudomonas
663	29	85.3	290	2	Q5BJN1_RAT	Q4s363_tetraodon n
664	29	85.3	291	1	PCTL_HUMAN	Q5bjnl_rattus norv
665	29	85.3	291	1	PCTL_MOUSE	Q9y365_homo sapien
666	29	85.3	295	2	Q4TNL4_9SPHN	Q9jmd3_mus musculu
667	29	85.3	296	1	KSGA_STAEQ	Q4tnl4_erythrobact
668	29	85.3	296	1	KSGA_STAEQ	Q8cqu5_s_dimethyla
669	29	85.3	296	2	Q4N4C6_THEPA	Q5hrr2_s_dimethyla
670	29	85.3	297	1	ZIPA_PSEPK	Q4n4c6_theileria p
671	29	85.3	298	2	Q5YPT6_NOCFA	Q88f24_pseudomonas
672	29	85.3	299	2	Q89E74_BRAJA	Q5ypt6_nocardia fa
673	29	85.3	301	2	Q9L526_VIBCH	Q89e74_bradyrhizob
674	29	85.3	301	2	Q4NHT3_9MICC	Q9l526_vibrio chol
675	29	85.3	301	2	Q9KKV6_VIBCH	Q4nht3_arthrobacte
676	29	85.3	305	2	Q86XV4_HUMAN	Q9kky6_vibrio chol
677	29	85.3	307	2	Q7W2V9_BORPA	Q86xv4_homo sapien
678	29	85.3	307	2	Q7WDV9_BORBR	Q7w2v9_bordetella
679	29	85.3	308	2	Q7MLJ5_PYLLI	Q7wdv9_bordetella
680	29	85.3	308	2	Q4NG45_9MICC	Q7mlj5_pylaiella l
681	29	85.3	308	2	Q6LG31_PHOPR	Q4ng45_arthrobacte
682	29	85.3	308	2	Q7Y5I0_9CAUD	Q6lg31_photobacter
683	29	85.3	309	2	Q7WBS3_BORPA	Q7y5i0_xanthomonas
684	29	85.3	309	2	Q7WPR6_BORBR	Q7wbs3_bordetella
685	29	85.3	311	1	GPDA_HELHP	Q7wpr6_bordetella
686	29	85.3	313	2	Q8XRE2_RALSO	Q7vil5_helicobacte
687	29	85.3	315	2	Q4HVD0_GIBZE	Q8xre2_ralstonia s
688	29	85.3	318	2	Q4GYV8_9TRYP	Q4hvd0_gibberella
689	29	85.3	318	2	Q4GYV8_9TRYP	Q4gyv8_trypanosoma

689	29	85.3	319	2	Q9RKM5_STRCO	Q9rkms	streptomyce
690	29	85.3	324	2	Q5L6A8_CHLAB	Q5l6a8	chlamydophi
691	29	85.3	324	2	Q5QX30_IDILO	Q5qx30	idiomarina
692	29	85.3	326	1	YGGR_ECOLI	P52052	escherichia
693	29	85.3	326	2	Q7NJV2_GLOVI	Q7njv2	gloeobacter
694	29	85.3	326	2	Q8JPA4_ROTTHU	Q8jpa4	human rotav
695	29	85.3	326	2	Q9WI74_9REOV	Q9wi74	human rotav
696	29	85.3	326	2	Q9WI76_9REOV	Q9wi76	human rotav
697	29	85.3	327	2	Q8RS23_STRVL	Q8rs23	streptomyce
698	29	85.3	327	2	Q9KPN2_VIBCH	Q9kpn2	vibrio chol
699	29	85.3	327	2	Q92RR1_RHIME	Q92rr1	rhizobium m
700	29	85.3	328	2	Q8WHN3_9ASTR	Q8whn3	xeranthemum
701	29	85.3	329	2	Q6A1K4_HELVI	Q6alk4	heliothis v
702	29	85.3	332	2	Q7S1I6_NEUCR	Q7si16	neurospora
703	29	85.3	337	2	Q83PN9_SHIFL	Q83pn9	shigella fl
704	29	85.3	337	2	Q8XDC2_ECO57	Q8xdc2	escherichia
705	29	85.3	338	1	LACD_LISIN	Q92eb7	listeria in
706	29	85.3	338	1	LACD_LISMF	Q723b1	listeria mo
707	29	85.3	338	1	LACD_LISMO	Q8y9i9	listeria mo
708	29	85.3	338	2	Q8R5T1_THETN	Q8rst1	thermoanaer
709	29	85.3	339	1	RIBD_MYCTU	P71677	m riboflavi
710	29	85.3	339	2	Q4YZR6_PLABE	Q4yzzr6	plasmodium
711	29	85.3	339	2	Q7U043_MYCBO	Q7u043	m probable
712	29	85.3	339	2	Q9CCP8_MYCLE	Q9ccp8	mycobacteri
713	29	85.3	341	2	Q741F5_MYCPA	Q741f5	mycobacteri
714	29	85.3	341	2	Q8XCU9_ECO57	Q8xcu9	escherichia
715	29	85.3	341	2	Q83JS3_SHIFL	Q83js3	shigella fl
716	29	85.3	343	2	Q6E76_AERPU	Q68e76	aeromonas p
717	29	85.3	343	2	Q6FB21_ACIAD	Q6fb21	acinetobact
718	29	85.3	344	2	Q9HN36_HALSA	Q9hn36	halobacteri
719	29	85.3	344	2	Q55829_9FLAV	Q55829	sokuluk vir
720	29	85.3	345	2	Q52LA1_MOUSE	Q52lal	mus musculu
721	29	85.3	347	2	Q9G5Y4_AGAIM	Q9g5y4	agama impal
722	29	85.3	350	1	MUTY_SALTY	Q05869	salmonella
723	29	85.3	350	2	Q57K05_SALCH	Q57k05	salmonella
724	29	85.3	350	2	Q823U0_SALTI	Q823u0	salmonella
725	29	85.3	350	2	Q5PMM2_SALPA	Q5pmm2	salmonella
726	29	85.3	350	2	Q4S273_TETNG	Q4s273	tetraodon n
727	29	85.3	351	2	Q5AZ00_EMENI	Q5az00	aspergillus
728	29	85.3	351	2	Q4HYD8_GIBZE	Q4hyd8	gibberella
729	29	85.3	351	2	Q95GI0_USEUD	Q95gi0	saxifragell
730	29	85.3	351	2	Q8FE29_ECOL6	Q8fe29	escherichia
731	29	85.3	352	1	ENGCI_VIBVU	Q8dcv7	vibrio vuln
732	29	85.3	352	1	ENGCI_VIBVY	Q7mgz6	vibrio vuln
733	29	85.3	352	1	MATK_SAXCE	Q33078	saxifraga c
734	29	85.3	353	2	Q66F15_YERPS	Q66f15	yersinia ps
735	29	85.3	353	2	Q8ZJ35_YERPE	Q8zj35	yersinia pe
736	29	85.3	355	2	Q75A37_ASHGO	Q75a37	ashbya goss
737	29	85.3	355	2	Q60DU2_ORYSA	Q60du2	oryza sativ
738	29	85.3	356	2	Q4LTP6_9BURK	Q4ltp6	burkholderi
739	29	85.3	356	2	Q67JN9_SYMTH	Q67jn9	symbiobacte
740	29	85.3	357	2	Q8GX19_ARATH	Q8gxi9	arabidopsis
741	29	85.3	357	2	Q9HY94_PSEAE	Q9hy94	pseudomonas
742	29	85.3	359	1	IDI2_METJA	Q58272	methanococc
743	29	85.3	360	2	Q8GLQ8_AERSA	Q8glq8	aeromonas s
744	29	85.3	361	2	Q62FF1_BURMA	Q62ff1	burkholderi
745	29	85.3	362	2	Q57U20_9TRYP	Q57u20	trypanosoma
746	29	85.3	362	2	Q60B22_METCA	Q60b22	methylococc
747	29	85.3	363	2	Q8TVS1_METKA	Q8tvsl	methanopyru
748	29	85.3	364	2	Q29447_ARCFU	Q29447	archaeoglob
749	29	85.3	364	2	Q67SV1_SYMTH	Q67sv1	symbiobacte
750	29	85.3	365	1	IDI2_METKA	Q8tx99	methanopyru
751	29	85.3	367	2	Q5KFI9_CRYNE	Q5kfi9	cryptococcu
752	29	85.3	369	2	Q9UZ22_PYRAB	Q9uz22	pyrococcus
753	29	85.3	369	2	Q5RL65_MOUSE	Q5rl65	mus musculu
754	29	85.3	370	2	Q8UJB9_AGRTS	Q8ujb9	agrobacteri
755	29	85.3	371	2	Q4J8E9_SULAC	Q4j8e9	sulfolobus
756	29	85.3	371	2	Q70SK6_9MYRT	Q70sk6	epilobium b
757	29	85.3	371	2	Q70SK8_9MYRT	Q70sk8	epilobium c
758	29	85.3	371	2	Q98HB1_RHILO	Q98hb1	rhizobium l
759	29	85.3	375	2	Q4NTI8_9DELT	Q4nti8	anaeromyxob
760	29	85.3	376	2	Q86SF7_HUMAN	Q86sf7	homo sapien
761	29	85.3	376	2	Q73TK1_MYCPA	Q73tk1	mycobacteri

29	85.3	378	1	MSRAB_VIBCH	Q9klx6	v peptide m
29	85.3	382	1	TGFB1_ONCMY	Q93449	oncorhynchu
29	85.3	382	2	Q7V7J6_PROMM	Q7v7j6	prochloroco
29	85.3	382	2	Q827H5_STRAW	Q827h5	streptomyce
29	85.3	383	1	BODG_RHILO	Q98kk0	rhizobium l
29	85.3	384	2	Q6PDD6_MOUSE	Q6pdd6	mus musculu
29	85.3	385	2	Q53CW4_9GAMA	Q53cw4	macaca fusc
29	85.3	386	2	Q86SI6_HUMAN	Q86si6	homo sapien
29	85.3	390	2	Q5YZ89_NOCFA	Q5yz89	nocardia fa
29	85.3	391	2	Q6UUH2_ORYSA	Q6uuh2	oryza sativ
29	85.3	392	2	Q54FP2_DICDI	Q54fp2	dictyosteli
29	85.3	392	2	Q6NY63_BRARE	Q6ny63	brachydanio
29	85.3	393	2	Q5L8Z1_BACFRN	Q5l8z1	bacterioides
29	85.3	393	2	Q64P54_BACFR	Q64p54	bacterioides
29	85.3	393	2	Q8A0W8_BACTN	Q8a0w8	bacterioides
29	85.3	395	2	Q5FQ25_GLUOX	Q5fq25	gluconobact
29	85.3	396	1	PORA_PYRFU	Q51804	pyrococcus
29	85.3	397	2	Q5V4T9_HALMA	Q5v4t9	haloarcula
29	85.3	399	2	Q9SSD0_ARATH	Q9ssd0	arabidopsis
29	85.3	405	1	IPSP_MOUSE	P70458	mus musculu
29	85.3	405	2	Q8N323_HUMAN	Q8n323	homo sapien
29	85.3	405	2	Q8BKC4_MOUSE	Q8bkc4	mus musculu
29	85.3	405	2	Q66JR2_MOUSE	Q66jr2	mus musculu
29	85.3	405	2	Q5BKQ8_MOUSE	Q5bkq8	mus musculu
29	85.3	405	2	Q8BU50_MOUSE	Q8bu50	m mus muscu
29	85.3	405	2	Q8BVN1_MOUSE	Q8bvn1	mus musculu
29	85.3	406	2	Q88292_RAT	Q88292	rattus norv
29	85.3	406	2	Q66HL5_RAT	Q66hl5	rattus norv
29	85.3	408	1	Y2112_PYRAE	Q8zvu9	pyrobaculum
29	85.3	408	2	Q92NV5_RHIME	Q92nv5	rhizobium m
29	85.3	409	2	Q6C525_YARLI	Q6c525	yarrowia li
29	85.3	410	2	Q4GYV2_9TRYP	Q4gyv2	trypanosoma
29	85.3	411	2	Q7VFN5_HELHP	Q7vfn5	helicobacte
29	85.3	413	1	DXR_NITEU	Q82u01	nitrosomona
29	85.3	414	1	Y4VG_RHISN	Q53215	rhizobium s
29	85.3	414	2	Q8UDV4_AGRTS	Q8udv4	agrobacteri
29	85.3	415	2	Q4HKI2_CAMLA	Q4hki2	campylobact
29	85.3	416	1	Y997_ARCFU	Q29265	archaeoglob
29	85.3	416	2	Q9EXQ4_CAMJE	Q9exq4	campylobact
29	85.3	416	2	Q4HHY9_CAMCO	Q4hhy9	campylobact
29	85.3	416	2	Q5HVA5_CAMJR	Q5hva5	campylobact
29	85.3	416	2	Q822P5_CHLCV	Q822p5	chlamydophi
29	85.3	416	2	Q9PPN1_CAMJE	Q9ppn1	campylobact
29	85.3	416	2	Q6ANX8_DESPS	Q6anx8	desulfotale
29	85.3	417	1	TRMU_YEAST	Q12093	saccharomyc
29	85.3	417	2	Q6IUF4_YEAST	Q6iuf4	saccharomyc
29	85.3	418	2	Q4HRQ8_CAMUP	Q4hrq8	campylobact
29	85.3	423	2	Q8X513_BCO57	Q8x513	escherichia
29	85.3	425	2	Q9NPN3_HUMAN	Q9npn3	homo sapien
29	85.3	425	2	Q83X64_STRRO	Q83x64	streptomyce
29	85.3	426	1	YJIN_ECOLI	P39385	escherichia
29	85.3	426	2	Q659Y0_ECOLI	Q659y0	escherichia
29	85.3	427	2	Q8G3V7_BIFLO	Q8g3v7	bifidobacte
29	85.3	427	2	Q8KG31_CHLTE	Q8kg31	chlorobium
29	85.3	427	2	Q86WN8_HUMAN	Q86wn8	homo sapien
29	85.3	433	2	Q4H3B0_CIOIN	Q4hb0	ciona intes
29	85.3	434	2	Q7MZJ8_PHOLL	Q7mzj8	photorhabdu
29	85.3	437	2	Q6JKF5_9NUCL	Q6jkf5	neodiprion
29	85.3	438	2	Q8FA91_ECOL6	Q8fa91	escherichia
29	85.3	439	2	Q6L5A2_ORYSA	Q6l5a2	oryza sativ
29	85.3	440	2	Q30538_PSEAE	Q30538	pseudomonas
29	85.3	440	2	Q8VWB7_9ACTO	Q8vwb7	streptomyce
29	85.3	440	2	Q9L555_9ACTO	Q9l555	streptomyce
29	85.3	440	2	Q9I330_PSEAE	Q9i330	pseudomonas
29	85.3	441	1	Y1340_HAEIN	P44165	haemophilus
29	85.3	445	2	Q70WP6_AERSA	Q70wp6	aeromonas s
29	85.3	449	1	Y942_SYNEL	Q8dkb7	synechococc
29	85.3	450	2	Q7SCC9_NEUCR	Q7scc9	neurospora
29	85.3	452	2	Q500J0_PSESY	Q500j0	pseudomonas
29	85.3	454	1	Y1462_HAEIN	P45217	haemophilus
29	85.3	454	2	Q4QKI6_HAEI8	Q4qki6	haemophilus
29	85.3	455	2	Q603F8_METCA	Q603f8	methylococc
29	85.3	455	2	Q8RH36_FUSNN	Q8rh36	fusobacteri

835 29 85.3 457 1 Y4BF RHISN
836 29 85.3 457 2 Q8XV7 RALSO
837 29 85.3 458 2 Q55QM6_CRYNE
838 29 85.3 459 2 Q86WN7_HUMAN
839 29 85.3 460 2 Q6MPF0_BDEBA
840 29 85.3 464 2 Q4J5N0_AZOVI
841 29 85.3 464 2 Q54204_STRCO
842 29 85.3 464 2 Q825R0_STRAW
843 29 85.3 465 1 MPPB_BLAEM
844 29 85.3 466 2 Q5NZR1_AZOSE
845 29 85.3 466 2 Q82JF1_STRAW
846 29 85.3 467 2 Q4Q160_LEIMA
847 29 85.3 467 2 Q5E6K2_VIBF1
848 29 85.3 470 2 Q7TM95_MOUSE
849 29 85.3 471 2 Q7W4T3_BORPA
850 29 85.3 472 2 Q8TZL5_PYRFU
851 29 85.3 473 2 Q7PKB4_ANOGA
852 29 85.3 473 2 Q8KL89_RHIET
853 29 85.3 474 2 Q7VU97_BORPE
854 29 85.3 474 2 Q7WGA9_BORBR
855 29 85.3 474 2 Q72AD1_DESVH
856 29 85.3 475 2 Q9KWD8_AGRRH
857 29 85.3 479 2 Q55IU1_CRYNE
858 29 85.3 479 2 Q5KCM0_CRYNE
859 29 85.3 482 2 Q9JPA7_RHOGE
860 29 85.3 483 2 Q5WND6_CAEBR
861 29 85.3 484 2 Q83YG8_RHOGE
862 29 85.3 485 2 Q5QD39_HUMAN
863 29 85.3 485 2 Q9I6T6_PSEAE
864 29 85.3 491 2 Q86WN6_HUMAN
865 29 85.3 492 2 Q86SJ3_HUMAN
866 29 85.3 497 2 Q4P263_USTMA
867 29 85.3 500 2 Q9H336_HUMAN
868 29 85.3 501 2 Q6N6Q2_RHOPA
869 29 85.3 502 2 Q98862_9ASTE
870 29 85.3 504 2 Q4UF69_THEAN
871 29 85.3 504 2 Q8GW59_ARATH
872 29 85.3 504 2 Q6X5V5_9ASTE
873 29 85.3 505 2 Q8N3F2_HUMAN
874 29 85.3 505 2 Q7UGI8_RHOBA
875 29 85.3 506 2 Q8MD40_9ASTE
876 29 85.3 506 2 Q8MD41_9ASTE
877 29 85.3 506 2 Q8MD43_9ASTE
878 29 85.3 507 2 Q4WL22_ASPFU
879 29 85.3 507 2 Q86WN9_HUMAN
880 29 85.3 510 2 Q6C469_YARLI
881 29 85.3 517 2 Q6BPB4_DEBHA
882 29 85.3 526 2 Q86WN5_HUMAN
883 29 85.3 530 2 Q5P026_AZOSE
884 29 85.3 531 2 Q87WE7_PSESM
885 29 85.3 533 2 Q53Y40_HUMAN
886 29 85.3 534 1 PDE9A_MOUSE
887 29 85.3 534 2 Q8QZV1_RAT
888 29 85.3 534 2 Q8BSU4_MOUSE
889 29 85.3 535 2 Q6EUF2_ORYSA
890 29 85.3 537 2 Q87LH6_NEUCR
891 29 85.3 537 2 O13369_ASCIM
892 29 85.3 540 1 BB61_RABIT
893 29 85.3 540 2 Q86WF0_HUMAN
894 29 85.3 541 2 Q8YL89_ANASP
895 29 85.3 542 2 Q5BDF1_EMENI
896 29 85.3 542 2 Q4WCK6_ASPFU
897 29 85.3 543 2 Q4P8Y4_USTMA
898 29 85.3 543 2 Q4IBN4_GIBZE
899 29 85.3 543 2 Q5GV08_XANOR
900 29 85.3 543 2 Q4V062_XANCP
901 29 85.3 543 2 Q8PDX0_XANCP
902 29 85.3 544 2 Q7S0Z6_NEUCR
903 29 85.3 547 2 Q7SFQ9_NEUCR
904 29 85.3 549 2 Q5Z9E4_ORYSA
905 29 85.3 549 2 Q92PP7_RHIME
906 29 85.3 551 2 Q9KXI6_STRCO
907 29 85.3 552 2 Q86WN4_HUMAN

P55373 rhizobium s
Q8xvv7 ralstonia s
Q55qm6 cryptococcu
Q86wn7 homo sapien
Q6mpf0 bdellovibri
Q4j5n0 azotobacter
Q54204 streptomyce
Q825r0 streptomyce
Q00302 blastocladi
Q5nzz1 azoarcus sp
Q82jfl streptomyce
Q4ql60 leishmania
Q5e6k2 vibrio fisc
Q7tm95 mus musculu
Q7w4t3 bordetella
Q8tzl5 pyrococcus
Q7pkb4 anopheles g
Q8kl89 rhizobium e
Q7vu97 bordetella
Q7wga9 bordetella
Q72ad1 desulfovibr
Q9kwd8 agrobacteri
Q55iu1 cryptococcu
Q5kcw0 cryptococcu
Q9jpa7 rhodocyclus
Q5wnd6 caenorhabdi
Q83yg8 rhodocyclus
Q5qd39 homo sapien
Q9i6tc6 pseudomonas
Q86wn6 homo sapien
Q86sj3 homo sapien
Q9h336 homo sapien
Q6n6q2 rhodopseudo
Q98862 petalonyx n
Q4uf69 theileria a
Q8gw59 arabidopsis
Q6x5v5 petalonyx l
Q8n3f2 homo sapien
Q7ugi8 rhodopirell
Q8md41 petalonyx t
Q8md43 petalonyx p
Q4wl22 aspergillus
Q6c469 yarrowia li
Q6bpb4 debaryomyce
Q86wn5 homo sapien
Q5p026 azoarcus sp
Q87we7 pseudomonas
Q53y40 homo sapien
Q70628 mus musculu
Q8qzv1 rattus norv
Q8bsu4 mus musculu
Q6euf2 oryza sativ
Q87lh6 neurospora
O13369 ascobolus i
Q05004 oryctolagus
Q86wp0 homo sapien
Q8yl89 anabaena sp
Q5bdf1 aspergillus
Q4wck6 aspergillus
Q4p8y4 ustilago ma
Q4ibn4 gibberella
Q5gv08 xanthomonas
Q4v062 xanthomonas
Q8pdx0 xanthomonas
Q7s0z6 neurospora
Q7sfq9 neurospora
Q5z9e4 oryza sativ
Q92pp7 rhizobium m
Q9kxi6 streptomyce
Q86wn4 homo sapien

557 2 Q4IIA7_GIBZE
558 2 Q69YK3_HUMAN
558 2 Q6FZX6_BARQU
558 2 Q6G3A3_BARHE
558 2 Q9CX72_MOUSE
558 2 Q8BM95_MOUSE
560 2 Q51JX3_MAGGR
562 2 Q8DMI6_MAGNEL
564 2 Q8QS39_9BETA
567 1 G6PIA_ORYSA
567 1 G6PI_MAIZE
567 2 Q86WN3_HUMAN
567 2 Q5CTF8_CRYPV
567 2 Q5CK89_CRYHO
567 2 Q8H3D6_ORYSA
567 2 Q652G2_ORYSA
567 2 Q8H8M6_ORYSA
568 2 Q4IPB0_GIBZE
568 2 Q94Z24_PYLLI
568 2 Q6AH33_LEIXX
570 1 UL49_HCMVA
570 2 Q6RXG9_HCMV
570 2 Q6SW82_HCMV
572 2 Q80FH8_9LUTE
572 2 Q91L47_9LUTE
572 2 Q91L51_9LUTE
572 2 Q91L56_9LUTE
572 2 Q91L59_9LUTE
572 2 Q9JH78_9LUTE
572 2 Q9QON8_9LUTE
574 2 Q72PW8_LEPIC
574 2 Q8F6B9_LEPIN
575 2 Q4PFD2_USTMA
576 2 Q7U2I8_MYCBO
576 2 P96408_MYCTU
580 2 Q89RH2_BRAJA
593 1 PDE9A_HUMAN
595 2 Q7SCV2_NEUCR
597 2 Q4WAU8_ASPFU
597 2 Q64753_ARATH
598 2 Q8I7N1_CAEEL
601 2 Q4SU34_TETNG
603 2 Q5E2P0_VIBF1
607 1 GLMS_CLOTE
607 2 Q59KX2_CANAL
607 2 Q51A19_ENTHI
608 2 Q4H636_9DEIO
611 2 Q5U483_XENLA
613 2 Q5YRV8_NOCFA
615 2 Q4IQ02_GIBZE
615 2 Q4S3B0_TETNG
617 2 Q8K297_MOUSE
617 2 Q6PGL1_MOUSE
619 1 SC6A3_MOUSE
619 1 SC6A3_RAT
619 2 Q9JJ41_MOUSE
619 2 Q9R0X6_MOUSE
619 2 Q8C009_MOUSE
620 1 SC6A3_HUMAN
620 1 SC6A3_MACFA
620 2 Q7JHP7_MACFA
620 2 Q7JHP8_MACMU
620 2 Q7JHP9_MACACA
620 2 Q9GJT5_SAIISC
620 2 Q9GJT7_SAIISC
621 2 Q4TRX1_9SPHN
622 2 Q8NBJ5_HUMAN
622 2 Q8NC64_HUMAN
622 2 Q57UX0_9TRYP
623 2 Q5LMM1_SILPO
625 2 Q6NVG7_MOUSE
625 2 Q8BWD4_MOUSE
629 2 Q7WVY1_VIBPS

Q4iia7 gibberella
Q69yk3 homo sapien
Q6fzx6 bartonella
Q6g3a3 bartonella
Q9cx72 mus musculu
Q8bm95 mus musculu
Q51jx3 magnaporthe
Q8dm16 synechococc
Q8qs39 pongine her
P42862 oryza sativ
P49105 zea mays (m
Q86wn3 homo sapien
Q5ctf8 cryptospori
Q5ck89 cryptospori
Q8h3d6 oryza sativ
Q652g2 oryza sativ
Q8h8m6 oryza sativ
Q4ipb0 gibberella
Q94z24 pylaiaella l
Q6ah33 leifsonia x
P16786 human cytom
Q6rxg9 human cytom
Q6sw82 human cytom
Q80fh8 sugarcane y
Q91l47 sugarcane y
Q91l51 sugarcane y
Q91l56 sugarcane y
Q91l59 sugarcane y
Q9jh78 sugarcane y
Q9qon8 sugarcane y
Q72pw8 leptospira
Q8f6b9 leptospira
Q4pfd2 ustilago ma
Q7u2i8 mycobacteri
P96408 mycobacteri
Q89rh2 bradyrhizob
Q89rh2 bradyrhizob
Q76083 homo sapien
Q7scv2 neurospora
Q4wau8 aspergillus
O64753 arabidopsis
Q8i7n1 caenorhabdi
Q4su34 tetraodon n
Q5e2p0 vibrio fisc
Q890u2 c glucosami
Q59kx2 candida alb
Q51a19 entamoeba h
Q4h636 deinococcus
Q5u483 xenopus lae
Q5yrv8 nocardia fa
Q4iq02 gibberella
Q4s3b0 tetraodon n
Q8k297 mus musculu
Q6pgl1 mus musculu
Q6l327 mus musculu
P23977 rattus norv
Q9jj41 mus musculu
Q9r0x6 mus musculu
Q8c009 mus musculu
Q01959 homo sapien
Q9gjt6 macaca fasc
Q7jhp7 macaca fasc
Q7jhp8 macaca mula
Q9gjt5 saimiri sci
Q9gjt7 saimiri sci
Q4trx1 erythrobaet
Q8nbj5 homo sapien
Q8nc64 homo sapien
Q57ux0 trypanosoma
Q5lmm1 silicibacte
Q6nvgt7 mus musculu
Q8bwd4 mus musculu
Q7wvy1 colwellia p

981 29 85.3 633 1 NUOL_MYCTU
982 29 85.3 633 2 Q7TX55_MYCBO
983 29 85.3 639 2 Q8TTH4_METAC
984 29 85.3 641 2 Q4QCK1_LEIMA
985 29 85.3 641 2 Q54MD1_DICDI
986 29 85.3 645 2 Q8C011_MOUSE
987 29 85.3 649 2 Q5STE6_CRYNE
988 29 85.3 649 2 Q5KIZ2_CRYNE
989 29 85.3 651 2 Q8T749_BRAFL
990 29 85.3 653 1 Y1364_MYCTU
991 29 85.3 653 2 Q7U068_MYCBO
992 29 85.3 654 2 Q5FQP2_GLUOX
993 29 85.3 657 2 Q61X34_CAEBR
994 29 85.3 662 2 Q5PAC8_ANAMM
995 29 85.3 663 2 Q7QF31_ANOGA
996 29 85.3 674 2 Q55T30_CRYNE
997 29 85.3 674 2 Q5KHG3_CRYNE
998 29 85.3 677 2 Q8BNT0_MOUSE
999 29 85.3 678 2 Q7UWE7_RHOBA
1000 29 85.3 682 2 Q4SHT0_TETNG

ALIGNMENTS

RESULT 1
Q8ZZ85_PYRAE
ID Q8ZZ85_PYRAE PRELIMINARY; PRT; 78 AA.
AC Q8ZZ85;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0388.
GN OrderedLocusNames=PAE0388;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009765; AAL62756.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 78 AA; 8694 MW; 382A6C0C24D130D3 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 17 TLREWL 22

RESULT 2
Q50WB5_ENTHI
ID Q50WB5_ENTHI PRELIMINARY; PRT; 78 AA.
AC Q50WB5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNAMES=173.t00016;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000543; EAL45891.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8632 MW; 3EA7ED258C3690F4 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 52 TLREWL 57

RESULT 3
Q9SHP3_ARATH
ID Q9SHP3_ARATH PRELIMINARY; PRT; 190 AA.
AC Q9SHP3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FIK23.24.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chao Q., Shinn P., Brooks S., Buehler E., Dunn P., Khan S., Kim C.,
RA Walker M., Brooks S., Altafi H., Araujo R., Conn L., Conway A.B.,
RA Gonzalez A., Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J.,
RA Liu S., Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V.,
RA Yu G., Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]

RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.


```
[4]
RN NUCLEOTIDE SEQUENCE.
RP Shinn P., Brooks S., Buehler E., Chao Q., Cheuk R., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007508; AAF24553.1; -; Genomic DNA.
SQ SEQUENCE 190 AA; 22057 MW; EA7455FDC49B0358 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 15 TLREWL 20

RESULT 4
Q4T3D3_TETNG
ID Q4T3D3_TETNG PRELIMINARY; PRT; 213 AA.
AC Q4T3D3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10097, whole genome shotgun sequence.
GN ORFNames=GSTENG0007901001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01010097; CAF92599.1; -; Genomic DNA.
SQ SEQUENCE 213 AA; 23088 MW; 23B8E3F0FF069D82 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 53 TLREWL 58
```

```
RESULT 5
Q7TNX8_RAT
ID Q7TNX8_RAT PRELIMINARY; PRT; 243 AA.
AC Q7TNX8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LRP16-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SD;
RA Guo J.H., Yu L.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF419856; AAP97291.1; -; mRNA.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR SMART; SM00506; Alpp; 1.
SQ SEQUENCE 243 AA; 27290 MW; DOF5933336EDD3BEF CRC64;

Query Match 100.0%; Score 34; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 204 TLREWL 209

RESULT 6
LRP16_RAT
ID LRP16_RAT STANDARD; PRT; 258 AA.
AC Q8K4G6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Protein LRP16 (Fragment).
GN Name=Lrp16;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Zhang X., Ip N.Y.;
RT "Rat mRNA sequence similar to LRP16 protein of humans.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- MISCELLANEOUS: Overexpression may promote MCF-7 cells
CC proliferation (By similarity).
CC -!- SIMILARITY: Contains 1 Alpp domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF404762; AAM45760.1; ALT INIT; mRNA.
DR Ensembl; ENSRNOG000000021174; Rattus norvegicus.
DR RGD; 628701; Lrp16.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR SMART; SM00506; Alpp; 1.
FT DOMAIN 103 217 Alpp.
FT NON TER 1 1
SQ SEQUENCE 258 AA; 28643 MW; FB430516A12D6B42 CRC64;
```

Query Match 100.0%; Score 34; DB 1; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 219 TLREWL 224

RESULT 7

KSGA_AZOSE STANDARD; PRT; 265 AA.
AC QSP7J1;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N',
DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High
DE level kasugamycin resistance protein ksgA) (Kasugamycin
DE dimethyltransferase).
GN Name=ksgA; OrderedLocusNames=AZOSEA05980; ORFNames=ebA1138;
OS Azocarcus sp. (strain EbN1).
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azocarcus.
OX NCBI_TaxID=76114;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15551059; DOI=10.1007/s00203-004-0742-9;
RA Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,
RA Reinhardt R.;
RT "The genome sequence of an anaerobic aromatic-degrading denitrifying
RT bacterium, strain EbN1.";
RL Arch. Microbiol. 183:27-36(2005).
CC -!- FUNCTION: Specifically dimethylates two adjacent adenosines in the
CC loop of a conserved hairpin near the 3'end of 16S rRNA in the 30S
CC particle. Its inactivation leads to kasugamycin resistance (By
CC similarity).
CC -!- SIMILARITY: Belongs to the rRNA adenine N-6-methyltransferase
CC family. KsgA subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; CR555306; CAI06720.1; -; Genomic_DNA.
DR HAMAP; MF_00607; -; 1.
DR InterPro; IPR011530; ksgA.
DR InterPro; IPR001737; RNA_meth_trans.
DR InterPro; IPR000051; SAM_Bd.
DR PANTHER; PTHR11727:SF6; ksgA; 1.
DR PANTHER; PTHR11727; RNA_meth_trans; 1.
DR Pfam; PF00398; RnaAD; 1.
DR SMART; SM00650; rADC; 1.
DR TIGRFAMs; TIGR00755; ksgA; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Complete proteome; Methyltransferase;
KW rRNA processing; Transferase.
SQ SEQUENCE 265 AA; 29742 MW; 239359142EAD5597 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 218 TLREWL 223

RESULT 8

Q87IG5_VIBPA PRELIMINARY; PRT; 277 AA.
ID Q87IG5_VIBPA PRELIMINARY;

AC Q87IG5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transcriptional regulator, LysR family.
GN OrderedLocusNames=VPA0641;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; BA000032; BAC61984.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTH_LYSR.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN_1.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Complete proteome.
SQ SEQUENCE 277 AA; 30745 MW; A7E211E9E24F23C2 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 129 TLREWL 134

RESULT 9

Q9KX21_OLICA PRELIMINARY; PRT; 296 AA.
ID Q9KX21_OLICA PRELIMINARY;
AC Q9KX21;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CoxK.
GN Name=coxK;
OS Oligotropha carboxidovorans (Pseudomonas carboxydovorans).
OG Plasmid pHCG3.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Oligotropha.
OX NCBI_TaxID=40137;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=OM5;
RX MEDLINE=95238294; PubMed=7721710;
RA Schuebel U., Kraut M., Moersdorf G., Meyer O.;
RT "Molecular characterization of the gene cluster coxMSL encoding the
RT molybdenum-containing carbon monoxide dehydrogenase of Oligotropha
RT carboxidovorans.";
RL J. Bacteriol. 177:2197-2203(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OM5;
RX MEDLINE=97464431; PubMed=9324252;
RA Santiago B., Meyer O.;

RT "Purification and molecular characterization of the H2 uptake
RT membrane-bound Nife-hydrogenase from the carboxidotrophic bacterium
RT Oligotropha carboxidovorans.";
RL J. Bacteriol. 179:6053-6060(1997).
RL [3]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RC Santiago B., Schuebel U., Egelseer C., Meyer O.;
RA "Sequence analysis, characterization and CO-specific transcription of
RT the cox gene cluster on the megaplasmid pHCg3 of Oligotropha
RT carboxidovorans.";
RL Gene 236:1157-1247(1999).
RL [4]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RC PubMed=14644498; DOI=10.1016/j.gene.2003.08.027;
RX Fuhrmann S., Ferner M., Jeffke T., Henne A., Gottschalk G., Meyer O.;
RA "Complete nucleotide sequence of the circular megaplasmid pHCg3 of
RT Oligotropha carboxidovorans: function in the chemolithoautotrophic
RT utilization of CO, H(2) and CO(2).";
RL Gene 322:67-75(2003).
RL [5]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RC Schuebel U.;
RA Submitted (JUL-1993) to the EMBL/GenBank/DBSJ databases.
RL [6]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RC Santiago B.;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.
RL [7]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RC Fuhrmann S.;
RA Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.
RL EMBL; X82447; CAB76251.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Plasmid.
SQ SEQUENCE 296 AA; 30979 MW; 174A75A5796CE9F2 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 74 TLREWL 79

RESULT 10
LRP16 HUMAN STANDARD; PRT; 325 AA.
ID LRP16_HUMAN Q9BQ69; Q9UH96;
AC Q9BQ69; Q9UH96;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Protein LRP16.
GN Name=LRP16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymphocyte;
RX PubMed=12578638;
RA Han W.-D., Yu L., Lou F.D., Wang Q.S., Zhao Y., Shi Z.J., Jin H.J.;
RT "The application of RACE technique to clone the full-length cDNA of a
RT novel leukemia associated gene LRP16.";

Zhongguo Shi Yan Xue Ye Xue Za Zhi 9:18-21(2001).
[2]
RL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RN TISSUE=Brain, and Eye;
RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [3]
RP INDUCTION.
RX MEDLINE=22675902; PubMed=12790785; DOI=10.1677/erc.0.0100217;
RA Han W.-D., Mu Y.-M., Lu X.-C., Xu Z.-M., Li X.-J., Yu L., Song H.-J.,
RA Li M., Lu J.-M., Zhao Y.-L., Pan C.-Y.;
RT "Up-regulation of LRP16 mRNA by 17beta-estradiol through activation of
RT estrogen receptor alpha (ERalpha), but not ERbeta, and promotion of
RT human breast cancer MCF-7 cell proliferation: a preliminary report.";
RL Endocr. Relat. Cancer 10:217-224(2003).
CC -!- INDUCTION: Overexpressed by estrogens in breast cancer MCF-7
CC cells, probably via an activation of nuclear receptors for
CC steroids (ESR1 and ESR2).
CC -!- MISCELLANEOUS: Overexpression may promote MCF-7 cells
CC proliferation.
CC -!- SIMILARITY: Contains 1 Alpp domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF202922; AAF15294.2; -; mRNA.
DR EMBL; BC000270; AAH00270.2; -; mRNA.
DR EMBL; BC003188; AAH03188.1; ALT INIT; mRNA.
DR EMBL; BC007297; AAH07297.1; -; mRNA.
DR EMBL; BC008316; AAH08316.1; -; mRNA.
DR HSSP; O28751; 1HJZ.
DR Ensembl; ENSG00000133315; Homo sapiens.
DR H-InvDB; HIX0019413; -.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR DOMAIN 170 284 Alpp.
SQ SEQUENCE 325 AA; 35505 MW; 82294BFC904FA4D0 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 286 TLREWL 291

RESULT 11
Q75PR0 PHYP
ID Q75PR0_PHYP PRELIMINARY; PRT; 391 AA.
AC Q75PR0;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Vascular plant one zinc finger protein.
GN Name=PpVOZ1;
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mitsuda N., Hisabori T., Takeyasu K., Sato M.H.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB164647; BAD12233.1; -; mRNA.
SQ SEQUENCE 391 AA; 42863 MW; 7B49223F8A834861 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 210 TLREWL 215

RESULT 12
Q7PSJ0 ANOQA PRELIMINARY; PRT; 473 AA.
AC Q7PSJ0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000012388 (Fragment).
GN ORFNames=ENSANGG00000009899;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008823; RAA05508.3; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1 1
FT NON_TER 473 473

SQ SEQUENCE 473 AA; 53729 MW; A994231087902F7C CRC64;

Query Match 100.0%; Score 34; DB 2; Length 473;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 280 TLREWL 285

RESULT 13
Q9SGQ0 ARATH PRELIMINARY; PRT; 486 AA.
AC Q9SGQ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE F3M18.4 (At1g28520) (Transcription factor AtVOZ1).
GN Name=AtVOZ1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Mitsuda N., Hisabori T., Takeyasu K., Sato M.H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010155; AAF16771.1; -; Genomic_DNA.
DR EMBL; BT020261; AAV84482.1; -; mRNA.
DR EMBL; AB125256; BAD17857.1; -; mRNA.
DR PIR; B86411; B86411.
SQ SEQUENCE 486 AA; 54079 MW; 9BA92951E2128858 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 311 TLREWL 316

RESULT 14
Q5BGN6 EMENI PRELIMINARY; PRT; 522 AA.
ID Q5BGN6_EMENI
AC Q5BGN6;

DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AN0294.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD0100006; EAA65700.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 522 AA; 58714 MW; 58DE8D4C9696F385 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 522;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 200 TLREWL 205

RESULT 15
Q6XPC4_PHYIN
ID Q6XPC4_PHYIN PRELIMINARY; PRT; 556 AA.
AC Q6XPC4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glucose-6-phosphate isomerase.
GN Name=GPI;
OS Phytophthora infestans (Potato late blight fungus).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4787;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22961941; PubMed=14599887; DOI=10.1016/S1087-1845(03)00107-5;
RA Ospina-Giraldo M.D., Jones R.W.;
RT "Characterization of the glucose-6-phosphate isomerase gene in
RT Phytophthora infestans reveals the presence of multiple alleles."
RL Fungal Genet. Biol. 40:197-206(2003).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-

CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AY220242; AAP51062.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
DR PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
SQ SEQUENCE 556 AA; 60840 MW; AACAC6C3EB68B752 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 221 TLREWL 226

RESULT 16
Q4HCY6_9DEIO
ID Q4HCY6_9DEIO PRELIMINARY; PRT; 557 AA.
AC Q4HCY6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=DgeODRAFT_1919;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000001; EAL83813.1; -; Genomic_DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 1 42 Potential.
SQ SEQUENCE 557 AA; 60854 MW; 1FC20B945E00DF6F CRC64;

Query Match 100.0%; Score 34; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 109 TLREWL 114

RESULT 17
Q5M7M1_XENTR
ID Q5M7M1_XENTR PRELIMINARY; PRT; 597 AA.
AC Q5M7M1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical LOC496849.
GN Name=LOC496849;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC088562; AAH88562.1; -; mRNA.
DR SNR; Q5M7M1; 302-590.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF00515; TPR 1; 4.
DR SMART; SM00028; TPR; 6.
DR PROSITE; PS50005; TPR; 5.
DR PROSITE; PS50293; TPR REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 597 AA; 67162 MW; BFF2D14AEA11CB87 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 410 TLREWL 415
|||||

RESULT 18
Q6FLE9_CANGA
ID Q6FLE9_CANGA PRELIMINARY; PRT; 637 AA.
AC Q6FLE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P53925 Saccharomyces cerevisiae YNL115c.
GN OrderedLocusNames=CAGL0L039389;
OS Candida glabrata (Yeast) (Torulopsis glabrata).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekaiia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380958; CAG61915.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR Pfam; PF00561; Abhydrolase_1; 1.
KW Complete proteome.
SQ SEQUENCE 637 AA; 73944 MW; D107B616C63CB755 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 261 TLREWL 266
|||||

RESULT 19
Q9VYJ0_DROME
ID Q9VYJ0_DROME PRELIMINARY; PRT; 963 AA.
AC Q9VYJ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG32648-PA.
GN Name=Pde9; ORFNames=CG32648;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426065; PubMed=12537568;
RX Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RL [3]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RL [4]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RL [5]
RN NUCLEOTIDE SEQUENCE.
RP Berkeley Drosophila Genome Project;
RG Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBBJ databases.
RN [6]
RN NUCLEOTIDE SEQUENCE.
RP FlyBase;
RG Submitted (MAR-2005) to the EMBL/GenBank/DBBJ databases.
RL -I- INTERACTION:
CC Q9VLM2:cg13384; NExp=1; IntAct=EBI-155878, EBI-109582;
CC EMBL; AE003490; AAF48205.2; -; Genomic_DNA.
DR HSSP; Q08499; 1OYN.
DR IntAct; Q9VYJ0; -.
DR Ensembl; CG32648; Drosophila melanogaster.

DR FlyBase; FBgn0052648; CG32648.
DR FlyBase; FBgn0052648; Pde9.
DR GO; GO:0004114; F:3,5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003607; Met_phos_hydro.
DR InterPro; IPR002073; PDease.
DR InterPro; IPR000217; Tubulin.
DR Pfam; PF00233; PDease_I; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR SMART; SM00471; HDC; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
DR PROSITE; PS00227; TUBULIN; UNKNOWN 1.
SQ SEQUENCE 963 AA; 103516 MW; B1987E0DFF6BD14F CRC64;

Query Match 100.0%; Score 34; DB 2; Length 963;
Best Local Similarity 100.0%; Pred.No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 93 TLREWL 98

RESULT 20
Q9UVC1 CLAFU
ID Q9UVC1 CLAFU PRELIMINARY; PRT; 1243 AA.
AC Q9UVC1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name-pol;
OS Cladosporium fulvum (Fulvia fulva).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
OX NCBI_TaxID=5499;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Simpson M.L., Butler M., Poulter R.T.M.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBBJ databases.
CC -I- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF051915; AAF21678.1; -; Genomic_DNA.
DR HSSP; P05205; 1KNA.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003682; F:chromatin binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00385; Chromo; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
DR SMART; SM00298; CHROMO; 1.
DR PROSITE; PS00598; CHROMO_1; 1.
DR PROSITE; PS50013; CHROMO_2; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW Nuclear protein; Polyprotein; Protease.
FT NON TER 1
SQ SEQUENCE 1243 AA; 143262 MW; C55999C5C80F58EE CRC64;

Query Match 100.0%; Score 34; DB 2; Length 1243;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches		6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1 TLREWL 6 									
Db	336 TLREWL 341									
RESULT 21										
Q9Z3T8	PSESX									
ID	Q9Z3T8_PSESX PRELIMINARY; PRT; 2066 AA.									
AC	Q9Z3T8;									
DT	01-MAY-1999 (TrEMBLrel. 10, Created)									
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)									
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)									
DE	Type I polyketide synthase.									
GN	Name=cfa7;									
OS	Pseudomonas syringae.									
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;									
OC	Pseudomonadaceae; Pseudomonas.									
OX	NCBI_TaxID=317;									
RN	[1]									
RP	NUCLEOTIDE SEQUENCE.									
RC	STRAIN=glycinea PG4180;									
RX	MEDLINE=99080036; PubMed=9860992; DOI=10.1073/pnas.95.26.15469;									
RA	Rangaswamy V., Jiralerspong S., Parry R., Bender C.L.;									
RT	"Biosynthesis of the Pseudomonas polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins.";									
RT	Proc. Natl. Acad. Sci. U.S.A. 95:15469-15474(1998).									
RRL	EMBL; AF098795; AAD03048.1; -; Genomic_DNA.									
DR	HSSP; P25715; 1MLA.									
DR	GO; GO:0048037; F:cofactor binding; IEA.									
DR	GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.									
DR	GO; GO:0016740; F:transferase activity; IEA.									
DR	GO; GO:0009058; P:biosynthesis; IEA.									
DR	GO; GO:0006633; P:fatty acid biosynthesis; IEA.									
DR	InterPro; IPR009081; ACP like.									
DR	InterPro; IPR001227; Ac transferase.									
DR	InterPro; IPR002198; ADH_short.									
DR	InterPro; IPR000794; Ketoacyl_synth.									
DR	InterPro; IPR006163; Phppanteth_bind.									
DR	InterPro; IPR001031; Thioesterase.									
DR	Pfam; PF00698; Acyl_transf_1; 1.									
DR	Pfam; PF00106; adh_short; 1.									
DR	Pfam; PF00109; ketoacyl-synt; 1.									
DR	Pfam; PF02801; Ketoacyl-synt_C; 1.									
DR	Pfam; PF00550; PP-binding; 1.									
DR	Pfam; PF00975; Thioesterase; 1.									
DR	PROSITE; PS50075; ACP_DOMAIN; 1.									
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.									
SQ	SEQUENCE 2066 AA; 221022 MW; EEP3F9DC8D20F751 CRC64;									
Query Match 100.0%; Score 34; DB 2; Length 2066;										
Best Local Similarity 100.0%; Pred. No. 1.9e+03;										
Matches		6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1 TLREWL 6 									
Db	1316 TLREWL 1321									
RESULT 22										
Q87W69	PSESX									
ID	Q87W69_PSESX PRELIMINARY; PRT; 2066 AA.									
AC	Q87W69;									
DT	01-JUN-2003 (TrEMBLrel. 24, Created)									
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)									
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)									
DE	Coronafacic acid polyketide synthetase II.									
GN	Name=cfa7; OrderedLocusNames=PSPT04687;									
OS	Pseudomonas syringae (pv. tomato).									
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;									
OC	Pseudomonadaceae; Pseudomonas.									
OX	NCBI_TaxID=323;									

RN	[1]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=DC3000;				
RX	MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;				
RA	Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,				
RA	Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,				
RA	Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,				
RA	Nelson W.C., Davidse T.M., Zafar N., Zhou L., Liu J., Yuan Q.,				
RA	Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,				
RA	Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,				
RA	Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,				
RA	Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,				
RA	Bender C.L., White O., Fraser C.M., Collmer A.;				
RT	"The complete genome sequence of the Arabidopsis and tomato pathogen				
RT	Pseudomonas syringae pv. tomato DC3000.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).				
DR	EMBL; AE016853; AAO58131.1; -; Genomic_DNA.				
DR	HSSP; P25715; 1MLA.				
DR	TIGR; PSPT04687; -.				
DR	GO; GO:0048037; F:cofactor binding; IEA.				
DR	GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.				
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.				
DR	GO; GO:0016740; F:transferase activity; IEA.				
DR	GO; GO:0009058; P:biosynthesis; IEA.				
DR	GO; GO:0006633; P:fatty acid biosynthesis; IEA.				
DR	InterPro; IPR009081; ACP_like.				
DR	InterPro; IPR001227; AC_transferase.				
DR	InterPro; IPR002198; ADH_short.				
DR	InterPro; IPR000794; Ketoacyl_synth.				
DR	InterPro; IPR006163; Phspanteth_bind.				
DR	InterPro; IPR001031; Thioesterase.				
DR	Pfam; PF00698; Acyl_transf_1; 1.				
DR	Pfam; PF00106; adh_short; 1.				
DR	Pfam; PF00109; ketoacyl-synt; 1.				
DR	Pfam; PF02801; Ketoacyl-synt_C; 1.				
DR	Pfam; PF00550; PP-binding; 1.				
DR	Pfam; PF00975; Thioesterase; 1.				
DR	PROSITE; PS50075; ACP_DOMAIN; 1.				
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.				
KW	Complete proteome.				
SQ	SEQUENCE 2066 AA; 221220 MW; E9C6D0274BB901DF CRC64;				
Query Match 100.0%; Score 34; DB 2; Length 2066;					
Best Local Similarity 100.0%; Pred. No. 1.9e+03;					
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1 TLREWL 6 				
Db	1316 TLREWL 1321				
RESULT 23					
Q6JD66	TOXGO				
ID	Q6JD66_TOXGO PRELIMINARY; PRT; 5072 AA.				
AC	Q6JD66;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	Eukaryotic initiation factor-2 alpha kinase-A.				
OS	Toxoplasma gondii.				
OC	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystida;				
OC	Toxoplasma.				
OX	NCBI_TaxID=5811;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RX	PubMed=14989696;				
RA	Sullivan W.J. Jr., Narasimhan J., Bhatti M.M., Wek R.C.;				
RT	"Parasite-specific eIF2 (eukaryotic initiation factor-2) kinase				
RT	required for stress-induced translation control.";				
RL	Biochem. J. 380:523-531(2004).				
CC	-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.				
CC	-!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.				
DR	EMBL; AY518936; AAS48463.1; -; mRNA.				

DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008627; GETHR.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF05671; GETHR; 10.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Initiation factor; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 5072 AA; 541699 MW; 00FD7BED9B848E58 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 5072;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 4707 TLREWL 4712

RESULT 24
Q6FEL5 ACIAD
ID Q6FEL5 ACIAD PRELIMINARY; PRT; 81 AA.
AC Q6FEL5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative acyl carrier protein (ACP).
GN OrderedLocusNames=ACIAD0569;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
CC -1- SIMILARITY: Contains 1 acyl carrier domain.
DR EMBL; CR543861; CAG67493.1; -; Genomic_DNA.
DR GO; GO:0000036; F:acyl carrier activity; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR003231; Acyl_carrier.
DR InterPro; IPR006163; Phsppanteth_bind.
DR Pfam; PF00550; PP-binding; 1.
DR ProDom; PD000887; Acyl_carrier; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KW Complete proteome; Phosphopantetheine.
SQ SEQUENCE 81 AA; 9345 MW; 5BB296B82DF58F60 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 81;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 10 TLREWM 15

RESULT 25

Q70SL4 9MYRT
ID Q70SL4 9MYRT PRELIMINARY; PRT; 134 AA.
AC Q70SL4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Hauya heydeana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Hauya.
OX NCBI_TaxID=225332;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ550743; CAD79575.1; -; Genomic_DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_isomerase.
DR Pfam; PF00342; PGI; 1.
KW Isomerase.
FT NON_TER 1
FT NON_TER 134
SQ SEQUENCE 134 AA; 14294 MW; E2F33E5CBC99A21B CRC64;

Query Match 94.1%; Score 32; DB 2; Length 134;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 96 TLREWI 101

RESULT 26
Q70SM0 9MYRT
ID Q70SM0 9MYRT PRELIMINARY; PRT; 134 AA.
AC Q70SM0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Circaea cordata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Circaeae.
OX NCBI_TaxID=13011;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ550737; CAD79569.1; -; Genomic_DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_isomerase.
DR Pfam; PF00342; PGI; 1.
KW Isomerase.
FT NON_TER 1
FT NON_TER 134
SQ SEQUENCE 134 AA; 14197 MW; 98EDD6882C1DB421 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 134;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 96 TLREWI 101
|||||:
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
|:|||||
Db 79 TMREWL 84
RESULT 27
Q70SL7_9MYRT PRELIMINARY; PRT; 144 AA.
AC Q70SL7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=PgiC;
OS Fuchsia cylindracea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Fuchsia.
OX NCBI_TaxID=225331;
RN [1]_NUCLEOTIDE SEQUENCE.
RP Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ550740; CAD79572.1; -; Genomic_DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15248 MW; E487F612A63A4CFE CRC64;
Query Match 94.1%; Score 32; DB 2; Length 144;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
|:|||||
Db 106 TLREWI 111
RESULT 28
Q656A5_ORYSA PRELIMINARY; PRT; 176 AA.
AC Q656A5;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Putative terminal flower 1.
GN Names=P0596H06.13;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT clone:P0596H06.",
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003620; BAD45362.1; -; Genomic_DNA.
DR Gramene; Q656A5; -;
DR InterPro; IPR001858; PBP.
DR InterPro; IPR008914; PEBP.
DR Pfam; PF01161; PBP; 1.
DR ProDom; PD004330; PBP; 1.
DR PROSITE; PS01220; PBP; 1.
SQ SEQUENCE 176 AA; 19207 MW; F79EC00390487391 CRC64;
Query Match 94.1%; Score 32; DB 2; Length 176;

Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
|:|||||
Db 79 TMREWL 84
RESULT 29
Q8D9L1_VIBVU PRELIMINARY; PRT; 195 AA.
AC Q8D9L1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=VV12586;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]_NUCLEOTIDE SEQUENCE.
RP STRAIN=CMCP6;
RC Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016805; AAO10937.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 195 AA; 21520 MW; 37B717876DF46C31 CRC64;
Query Match 94.1%; Score 32; DB 2; Length 195;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
|:|||||
Db 124 TLREWM 129
RESULT 30
Q6N1X5_RHOPA PRELIMINARY; PRT; 209 AA.
AC Q6N1X5;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RPA4277;
OS Rhodopsseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopsseudomonas.
OX NCBI_TaxID=1076;
RN [1]_NUCLEOTIDE SEQUENCE.
RP STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopsseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572606; CAE29718.1; -; Genomic_DNA.
DR GO; GO:0007076; P:mitotic chromosome condensation; IEA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84DB040EE CRC64;
Query Match 94.1%; Score 32; DB 2; Length 209;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
Db 109 TLREWL 114

RESULT 31
Q70SK1_9MYRT PRELIMINARY; PRT; 269 AA.
AC Q70SK1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Lopezia grandiflora.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Lopezia.
OX NCBI_TaxID=225335;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550756; CAD79588.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29125 MW; 0F67421013909157 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 269;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
Db 96 TLREWI 101

RESULT 32
Q70SK3_9MYRT PRELIMINARY; PRT; 281 AA.
AC Q70SK3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiCb;
OS Megacorax graciellanus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Megacorax.
OX NCBI_TaxID=203949;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550754; CAD79586.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 281
SQ SEQUENCE 281 AA; 30422 MW; 68BB058315299731 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
Db 96 TLREWI 101

RESULT 33
Q8RW74_CLAEP PRELIMINARY; PRT; 307 AA.
AC Q8RW74;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC2;
OS Clarkia epilobioides (Willow-herb clarkia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=145959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LDG 9605;
RX MEDLINE=22033959; PubMed=12038528;
RA Ford V.S., Gottlieb L.D.;
RT "Single mutations silence pgiC2 genes in two very recent allotetraploid species of Clarkia."
RL Evolution 56:699-707(2002).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ311754; CAC84516.1; -; Genomic DNA.
DR HSSP; P08059; 1GZD.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 307
SQ SEQUENCE 307 AA; 33553 MW; A62C334EBF309B25 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 307;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||

Db 156 TLREWI 161

RESULT 34

TOC34 ARATH STANDARD; PRT; 313 AA.

AC Q38906; Q8L9G5; Q9FF74; Q9GDD2;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Translocase of chloroplast 34 (34 kDa chloroplast outer envelope protein) (GTP-binding protein OEP34) (AtToc34).

DE Name=toc34; Synonyms=oe34; OrderedLocusNames=At5g05000;

GN ORFNames=MUG13.14;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Columbia;

RA Chen L.-J., Li H.-M.;

RT "Arabidopsis homologue of OEP34, a component of the pea chloroplast protein import apparatus.";

RL (er) Plant Gene Register PGR95-140.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Columbia;

RX MEDLINE=20452907; PubMed=10998188;

RA Gutensohn M., Schulz B.I., Nicolay P., Fluegge U.I.;

RT "Functional analysis of the two Arabidopsis homologues of Toc34, a component of the chloroplast protein import apparatus.";

RL Plant J. 23:771-783(2000).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=cv. Columbia;

RX MEDLINE=97471969; PubMed=9330910;

RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M., Miyajima N., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones.";

RL DNA Res. 4:215-230(1997).

RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R., Feldmann K.A.;

RT "Full-length cDNA from Arabidopsis thaliana.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SUBCELLULAR LOCATION.

RX PubMed=12775770; DOI=10.1074/mcp.M300030-MCP200;

RA Ferro M., Salvi D., Brugiére S., Miras S., Kowalski S., Louwagie M., Garin J., Joyard J., Rolland N.;

RT "Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana.";

RL Mol. Cell. Proteomics 2:325-345(2003).

CC -!- FUNCTION: Involved in protein import into chloroplasts. Seems to recognize chloroplast-destined precursor proteins and regulate their presentation to the translocation channel through GTP hydrolysis.

CC -!- SUBUNIT: Interacts with Toc75.

CC -!- SUBCELLULAR LOCATION: Chloroplast; outer membrane.

CC -!- SIMILARITY: Belongs to the TOC34 family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

CC EMBL; U43377; AAD09203.1; -; mRNA.

DR EMBL; AJ132696; CAC17699.1; -; mRNA.

DR EMBL; AB005245; BAB11522.1; -; Genomic_DNA.

DR EMBL; AY088447; AAM65983.1; -; mRNA.

DR InterPro; IPR006703; AIG1.

DR InterPro; IPR005688; Toc34.

DR Pfam; PF04548; AIG1; 1.

DR TIGRFAMs; TIGR00991; 3a0901s02IAP34; 1.

KW Chloroplast; GTP-binding; Membrane; Nucleotide-binding;

KW Outer membrane; Protein transport; Transport.

FT NP_BIND 45 52 GTP (Potential).

FT NP_BIND 92 96 GTP (Potential).

FT NP_BIND 154 157 GTP (Potential).

FT CONFLICT 44 44 M -> I (in Ref. 4).

FT CONFLICT 47 47 G -> S (in Ref. 4).

FT CONFLICT 299 299 L -> P (in Ref. 3 and 4).

SQ SEQUENCE 313 AA; 34723 MW; 1C90E435B2EE1F2D CRC64;

Query Match 94.1%; Score 32; DB 1; Length 313;

Best Local Similarity 83.3%; Pred.No. 6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 6 TLREWI 11

RESULT 35

Q5V5F2 HALMA

ID Q5V5F2_HALMA PRELIMINARY; PRT; 328 AA.

AC Q5V5F2;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)

DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Putative transporter.

GN OrderedLocusNames=rrnAC0187;

OS Haloarcula marismortui (Halo bacterium marismortui).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

OC Halobacteriaceae; Haloarcula.

OX NCBI_TaxID=2238;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 43049;

RX PubMed=15520287; DOI=10.1101/gr.2700304;

RA Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G., Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P., Date S.V., Marcotte E., Hood L., Ng W.V.;

RT "Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea.";

RL Genome Res. 14:2221-2234(2004).

DR EMBL; AY596297; AAV45250.1; -; Genomic_DNA.

KW Complete proteome.

SQ SEQUENCE 328 AA; 35627 MW; 4DC488FB91107046 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 328;

Best Local Similarity 83.3%; Pred.No. 6.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 293 TLREWI 298

RESULT 36

Q6MB42 PARUW

ID Q6MB42_PARUW PRELIMINARY; PRT; 342 AA.

AC Q6MB42;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=pcl483;

OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).

OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.

RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006619; BAA22037.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER_target_S.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
FT NON_TER 357
SQ SEQUENCE 357 AA; 39029 MW; 046D30C542200FF4 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 153 TLREWI 158
|||||:

RESULT 40
O22095_9LILI PRELIMINARY; PRT; 357 AA.
AC O22095;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea tenuipes.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98172211; PubMed=95111221;
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006620; BAA22038.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER_target_S.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
FT NON_TER 357
SQ SEQUENCE 357 AA; 39029 MW; 046D30C542200FF4 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 153 TLREWI 158
|||||:

RESULT 40
O22095_9LILI PRELIMINARY; PRT; 357 AA.
AC O22095;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea tenuipes.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98172211; PubMed=95111221;
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006620; BAA22038.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER_target_S.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.

DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
FT NON_TER 357
SQ SEQUENCE 357 AA; 39078 MW; DE9C1AE79030E6AD CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 153 TLREWI 158
|||||:

RESULT 41
O24644_9LILI PRELIMINARY; PRT; 357 AA.
AC O24644;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea gracillima.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98172211; PubMed=95111221;
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006616; BAA22034.1; -; mRNA.
DR EMBL; AB006615; BAA22033.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER_target_S.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
FT NON_TER 357
SQ SEQUENCE 357 AA; 39021 MW; 08DC40108A3B165B CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 153 TLREWI 158
|||||:

RESULT 42
Q4SIN4_TETNG
ID Q4SIN4_TETNG PRELIMINARY; PRT; 368 AA.
AC Q4SIN4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)

Db 96 TLREWI 101

RESULT 45

Q70SK5 9MYRT

ID Q70SK5 9MYRT PRELIMINARY; PRT; 373 AA.

AC Q70SK5

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).

GN Name=pgiC2;

OS Epilobium brachycarpum.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; Myrtales; Onagraceae; Epilobium.

OX NCBI_TaxID=29754;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Gottlieb L.D., Ford V.S.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.

CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the GPI family.

DR EMBL; AJ550752; CAD79584.1; -; Genomic DNA.

DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.

DR GO; GO:0016853; F:isomerase activity; IEA.

DR GO; GO:0006094; P:gluconeogenesis; IEA.

DR GO; GO:0006096; P:glycolysis; IEA.

DR InterPro; IPR001672; G6P_Isomerase.

DR Pfam; PF00342; PGI; 1.

DR PRINTS; PR00662; G6PISOMERASE.

DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.

KW Gluconeogenesis; Glycolysis; Isomerase.

FT NON_TER 1

FT NON_TER 373

FT NON_TER 373

SQ SEQUENCE 373 AA; 40728 MW; 2425C5F1F6D1B4FE CRC64;

Query Match 94.1%; Score 32; DB 2; Length 373;

Best Local Similarity 83.3%; Pred. No. 7.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6

Db 96 TLREWI 101

RESULT 46

Q8L6V2 9MYRT

ID Q8L6V2 9MYRT PRELIMINARY; PRT; 380 AA.

AC Q8L6V2;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).

GN Name=pgiC2b;

OS Clarkia dudleyana.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; Myrtales; Onagraceae; Clarkia.

OX NCBI_TaxID=188038;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LDG 8624a;

RA Ford V.S., Gottlieb L.D.;

RL "Reassessment of phylogenetic relationships in Clarkia sect. Sympherica.";

RL Am. J. Bot. 90:284-292(2003).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LDG 8624a;

RX MEDLINE=22189021; PubMed=12200488;

RA Gottlieb L.D., Ford V.S.;

RT "The 5' leader of plant PgiC has an intron: the leader shows both the loss and maintenance of constraints compared with introns and exons in the coding region.";

RT Mol. Biol. Evol. 19:1613-1623(2002).

CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.

CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the GPI family.

DR EMBL; AJ437277; CAD24790.1; -; Genomic DNA.

DR HSSP; P08059; IGZD.

DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.

DR GO; GO:0016853; F:isomerase activity; IEA.

RC STRAIN=LDG 8624a;

RX MEDLINE=22189021; PubMed=12200488;

RA Gottlieb L.D., Ford V.S.;

RT "The 5' leader of plant PgiC has an intron: the leader shows both the loss and maintenance of constraints compared with introns and exons in the coding region.";

RL Mol. Biol. Evol. 19:1613-1623(2002).

CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.

CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the GPI family.

DR EMBL; AJ437278; CAD24791.1; -; Genomic DNA.

DR HSSP; P08059; IGZD.

DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.

DR GO; GO:0016853; F:isomerase activity; IEA.

DR GO; GO:0006094; P:gluconeogenesis; IEA.

DR GO; GO:0006096; P:glycolysis; IEA.

DR InterPro; IPR001672; G6P_Isomerase.

DR Pfam; PF00342; PGI; 1.

DR PRINTS; PR00662; G6PISOMERASE.

DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.

KW Gluconeogenesis; Glycolysis; Isomerase.

FT NON_TER 380

FT NON_TER 380

SQ SEQUENCE 380 AA; 41687 MW; 6871252AAFC5B066 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 380;

Best Local Similarity 83.3%; Pred. No. 7.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6

Db 226 TLREWI 231

RESULT 47

Q8L6V3 9MYRT

ID Q8L6V3 9MYRT PRELIMINARY; PRT; 380 AA.

AC Q8L6V3;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).

GN Name=pgiC2a;

OS Clarkia dudleyana.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; Myrtales; Onagraceae; Clarkia.

OX NCBI_TaxID=188038;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LDG 8624a;

RA Ford V.S., Gottlieb L.D.;

RT "Reassessment of phylogenetic relationships in Clarkia sect. Sympherica.";

RL Am. J. Bot. 90:284-292(2003).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LDG 8624a;

RX MEDLINE=22189021; PubMed=12200488;

RA Gottlieb L.D., Ford V.S.;

RT "The 5' leader of plant PgiC has an intron: the leader shows both the loss and maintenance of constraints compared with introns and exons in the coding region.";

RT Mol. Biol. Evol. 19:1613-1623(2002).

CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-phosphate.

CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the GPI family.

DR EMBL; AJ437277; CAD24790.1; -; Genomic DNA.

DR HSSP; P08059; IGZD.

DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.

DR GO; GO:0016853; F:isomerase activity; IEA.

DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 380 380
SQ SEQUENCE 380 AA; 41631 MW; AC59DFE13021CE7F CRC64;

Query Match 94.1%; Score 32; DB 2; Length 380;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 226 TLREWI 231
|||||:
226 TLREWI 231

RESULT 48
Q8L6V4 9MYRT
ID Q8L6V4 9MYRT PRELIMINARY; PRT; 380 AA.
AC Q8L6V4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC1;
OS Clarkia dudleyana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=188038;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LDG 8624a; PubMed=12200488;
RX MEDLINE=22189021; PubMed=12200488;
RA Gottlieb L.D., Ford V.S.;
RT "The 5' leader of plant PgiC has an intron: the leader shows both the
RT loss and maintenance of constraints compared with introns and exons in
RT the coding region.";
RL Mol. Biol. Evol. 19:1613-1623(2002).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ437276; CAD24789.1; -; Genomic_DNA.
DR HSSP; Q9N1E2; 1HM5.
DR GO; GO:0004347; F:Glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:Isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 380 380
SQ SEQUENCE 380 AA; 41659 MW; F6E18DF30212F586 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 380;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 226 TLREWI 231
|||||:
226 TLREWI 231

RESULT 49
Q8RW61 CLAUN
ID Q8RW61 CLAUN PRELIMINARY; PRT; 380 AA.
AC Q8RW61;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC2;
OS Clarkia unguiculata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=3937;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LDG-8829-8d-3A;
RX MEDLINE=22033959; PubMed=12038528;
RA Ford V.S., Gottlieb L.D.;
RT "Single mutations silence PgiC2 genes in two very recent
RT allotetraploid species of Clarkia.";
RL Evolution 56:699-707(2002).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ312370; CAC86124.1; -; Genomic_DNA.
DR HSSP; Q9N1E2; 1HM5.
DR GO; GO:0004347; F:Glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:Isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 380 380
SQ SEQUENCE 380 AA; 41762 MW; CCC72616F23BFB7F CRC64;

Query Match 94.1%; Score 32; DB 2; Length 380;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 226 TLREWI 231
|||||:
226 TLREWI 231

RESULT 50
Q70SL1 LUDPE
ID Q70SL1 LUDPE PRELIMINARY; PRT; 381 AA.
AC Q70SL1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Ludwigia peploides (Floating primrose-willow).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Ludwigia.
OX NCBI_TaxID=13119;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-

CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550746; CAD79578.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1 1
FT NON_TER 381 381
SQ SEQUENCE 381 AA; 41846 MW; F54CC9FC3161EDC2 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 381;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 106 TLREWI 111

Search completed: May 12, 2006, 10:51:28
Job time : 74.4615 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 72.6496 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-35
Perfect score: 56
Sequence: 1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	2	AAY22355 TPO recep
2	56	100.0	10	3	AAB16979 TPO-mimet
3	56	100.0	10	5	ABB72865 TPO mimet
4	56	100.0	10	7	ADJ73016 TPO mimet
5	56	100.0	10	8	ADJ52651 CH1 delet
6	56	100.0	10	8	ADJ51612 CH1 delet
7	41	73.2	780	8	ADN20123 Bacterial
8	41	73.2	787	6	ADA36768 Acinetoba
9	40	71.4	96	4	AAM92258 Human dig
10	40	71.4	353	9	ABM91714 M. xanthu
11	40	71.4	363	8	ADN21256 Bacterial
12	39	69.6	95	3	AAB54291 Human pan
13	39	69.6	558	4	AAB47456 D-aminoac
14	38	67.9	358	7	ABO62198 Klebsiell
15	38	67.9	536	2	AAW64365 Mycobacte
16	38	67.9	536	2	AAW81732 M. tuberc
17	38	67.9	536	2	AAY39019 M. tuberc
18	38	67.9	536	2	AAY39162 M. tuberc
19	37	66.1	103	4	AAU54158 Propionib
20	37	66.1	103	6	ABM50677 Propionib
21	37	66.1	206	7	ABO80351 Pseudomon
22	37	66.1	271	8	ABO84669 Mouse can
23	37	66.1	297	4	AAG92549 C glutami
24	37	66.1	423	8	ADT57937 Plant pol

98	35	62.5	500	6	ABP57627	Abp57627 S. muraya	171	34	60.7	230	8	ADX87466	Adx87466 Plant ful
99	35	62.5	507	8	ADT60952	Adt60952 Plant pol	172	34	60.7	236	4	AAU63835	Aau63835 Propionib
100	35	62.5	520	4	ABU52637	Abu52637 Human tra	173	34	60.7	236	6	ABM60354	Abm60354 Propionib
101	35	62.5	520	4	AAG67137	Aag67137 Amino aci	174	34	60.7	236	8	ADX66559	Adx66559 Plant ful
102	35	62.5	520	8	ADS23199	Ads23199 Bacterial	175	34	60.7	264	8	ADN21928	Adn21928 Bacterial
103	35	62.5	547	7	ABO77986	AbO77986 Pseudomon	176	34	60.7	293	9	ABM96423	Abm96423 M. xanthu
104	35	62.5	607	5	AAE22155	Aae22155 Human TRN	177	34	60.7	296	7	ADG10467	Adg10467 Human STA
105	35	62.5	607	6	ADA54308	Ada54308 Human pro	178	34	60.7	303	7	ABM87579	Abm87579 Rice abio
106	35	62.5	607	7	ADC35389	Adc35389 Human N-a	179	34	60.7	318	7	ABO78573	AbO78573 Pseudomon
107	35	62.5	607	7	ADE47686	Ade47686 Human NOV	180	34	60.7	343	7	ADD43962	Add43962 Chlamydia
108	35	62.5	607	7	ADE47688	Ade47688 Human NOV	181	34	60.7	343	9	AEA19254	Aea19254 Chlamydia
109	35	62.5	607	7	ADE47690	Ade47690 Human NOV	182	34	60.7	345	8	ADX80524	Adx80524 Plant ful
110	35	62.5	607	8	ADJ78958	Adj78958 Human NOV	183	34	60.7	351	8	ADN21155	Adn21155 Bacterial
111	35	62.5	607	8	ADJ78956	Adj78956 Human NOV	184	34	60.7	375	2	AAW20731	Aaw20731 H. pylori
112	35	62.5	607	8	ADJ78960	Adj78960 Human NOV	185	34	60.7	380	4	ABB06901	Abb06901 Micromono
113	35	62.5	620	7	ABO72087	AbO72087 Pseudomon	186	34	60.7	380	6	ABP99317	Abp99317 Orthosomy
114	35	62.5	944	7	ADE08076	Ade08076 Novel pro	187	34	60.7	380	8	ADT59120	Adt59120 Plant pol
115	35	62.5	944	9	ADU40446	Adu40446 Novel hum	188	34	60.7	413	8	ADX68804	Adx68804 Plant ful
116	35	62.5	948	4	ABB63850	Abb63850 Drosophil	189	34	60.7	437	2	AAy37663	Aay37663 Amino aci
117	35	62.5	971	4	ABG24290	Abg24290 Novel hum	190	34	60.7	444	8	ADX72887	Adx72887 Plant ful
118	35	62.5	1054	9	AEB69202	Aeb69202 Human mod	191	34	60.7	454	6	ABU45271	Abu45271 Protein e
119	35	62.5	1117	8	ADW64308	Adw64308 High-temp	192	34	60.7	457	6	ABU28329	Abu28329 Protein e
120	35	62.5	1234	6	ABM68967	Abm68967 Photorhab	193	34	60.7	457	6	ABU32229	Abu32229 Protein e
121	35	62.5	1515	6	ABU19676	Abu19676 Protein e	194	34	60.7	457	6	ABU15420	Abu15420 Protein e
122	35	62.5	2600	6	ABU22788	Abu22788 Protein e	195	34	60.7	457	6	ABU50197	Abu50197 Protein e
123	34.5	61.6	275	3	AAy87221	Aay87221 Human sec	196	34	60.7	457	8	ADS45219	Ads45219 Bacteriel
124	34.5	61.6	275	4	AAE06203	Aae06203 Human gen	197	34	60.7	458	6	ABU48015	Abu48015 Protein e
125	34.5	61.6	275	5	ABG34027	Abg34027 Human sec	198	34	60.7	458	6	ABU47154	Abu47154 Protein e
126	34.5	61.6	319	3	AAy87223	Aay87223 Human sec	199	34	60.7	461	7	ABO65735	AbO65735 Klebsiell
127	34.5	61.6	319	4	AAE06205	Aae06205 Human gen	200	34	60.7	488	6	ABU30709	Abu30709 Protein e
128	34.5	61.6	319	5	ABG34029	Abg34029 Human sec	201	34	60.7	490	5	ABP65879	Abp65879 Bifidobac
129	34.5	61.6	319	5	ABG33955	Abg33955 Human sec	202	34	60.7	548	2	AAR22197	Aar22197 Sequence
130	34.5	61.6	319	5	ABG33955	Abg33955 Human sec	203	34	60.7	575	8	ADY04932	Ady04932 Plant ful
131	34.5	61.6	319	6	ADA57361	Ada57361 Human sec	204	34	60.7	587	7	ABM90283	Abm90283 Rice abio
132	34.5	61.6	319	6	ADA41242	Ada41242 Human sec	205	34	60.7	609	7	ABO83801	AbO83801 Pseudomon
133	34.5	61.6	319	6	ABR48003	Abr48003 Human sec	206	34	60.7	660	2	AAR70013	Aar70013 Tick carb
134	34.5	61.6	320	3	AAy87155	Aay87155 Human sec	207	34	60.7	801	8	ADX88989	Adx88989 Plant ful
135	34.5	61.6	556	7	ADD19311	Add19311 Human sec	208	34	60.7	859	8	ADQ37037	Adq37037 Cell prol
136	34	60.7	9	7	ADJ73026	Adj73026 TPO mimet	209	34	60.7	927	7	ABO64572	AbO64572 Klebsiell
137	34	60.7	9	8	ADJ51622	Adj51622 CH1 delet	210	34	60.7	977	3	AAG51199	Aag51199 Arabidops
138	34	60.7	11	2	AAy22367	Aay22367 TPO recep	211	34	60.7	992	3	AAG51198	Aag51198 Arabidops
139	34	60.7	11	2	AAy22365	Aay22365 TPO recep	212	34	60.7	1051	4	ABB60157	Abb60157 Drosophil
140	34	60.7	11	2	AAy22370	Aay22370 TPO recep	213	34	60.7	1115	3	AAG51197	Aag51197 Arabidops
141	34	60.7	11	3	AAB16989	Aab16989 TPO-mimet	214	34	60.7	1214	4	AAU81522	Aau81522 C.elegans
142	34	60.7	11	3	AAB16986	Aab16986 TPO-mimet	215	34	60.7	1221	7	ADM25443	Adm25443 Hyperther
143	34	60.7	11	3	AAB16988	Aab16988 TPO-mimet	216	34	60.7	1529	3	AAy96744	Aay96744 A. terreu
144	34	60.7	11	5	ABB72872	Abb72872 TPO mimet	217	34	60.7	1711	8	ADN46550	Adn46550 Thermococ
145	34	60.7	11	5	ABB72875	Abb72875 TPO mimet	218	34	60.7	1711	8	ADN46550	Adn46550 Thermococ
146	34	60.7	11	5	ABB72874	Abb72874 TPO mimet	219	34	60.7	4437	6	ABU11384	Abu11384 Protein e
147	34	60.7	11	7	ADJ73023	Adj73023 TPO mimet	220	34	60.7	19938	6	ABP76680	Abp76680 Streptomy
148	34	60.7	11	7	ADJ73025	Adj73025 TPO mimet	221	33.5	59.8	633	9	ABM96728	Abm96728 M. xanthu
149	34	60.7	11	8	ADJ52658	Adj52658 CH1 delet	222	33	58.9	7	5	AAO18380	Aao18380 Lead stru
150	34	60.7	11	8	ADJ52661	Adj52661 CH1 delet	223	33	58.9	15	2	AAy22406	Aay22406 TPO recep
151	34	60.7	11	8	ADJ52660	Adj52660 CH1 delet	224	33	58.9	54	9	AEb18127	Aeb18127 Lactobaci
152	34	60.7	11	8	ADJ51619	Adj51619 CH1 delet	225	33	58.9	57	4	AAU57120	Aau57120 Propionib
153	34	60.7	11	8	ADJ51621	Adj51621 CH1 delet	226	33	58.9	57	6	ABM53639	Abm53639 Propionib
154	34	60.7	13	2	AAy22407	Aay22407 TPO recep	227	33	58.9	85	8	ADY25087	Ady25087 Plant ful
155	34	60.7	50	2	AAy07987	Aay07987 Human sec	228	33	58.9	103	4	AAU43490	Aau43490 Propionib
156	34	60.7	63	5	ABP00760	Abp00760 Human ORF	229	33	58.9	103	6	ABM40009	Abm40009 Propionib
157	34	60.7	73	4	AAM92259	Aam92259 Human dig	230	33	58.9	122	4	AAU59364	Aau59364 Propionib
158	34	60.7	76	2	AAy74245	Aay74245 Human pro	231	33	58.9	122	6	ABM55883	Abm55883 Propionib
159	34	60.7	80	4	AAM82693	Aam82693 Human imm	232	33	58.9	134	4	AAU52483	Aau52483 Propionib
160	34	60.7	124	7	ABO65668	AbO65668 Klebsiell	233	33	58.9	134	6	ABM49002	Abm49002 Propionib
161	34	60.7	148	7	ABO84012	AbO84012 Pseudomon	234	33	58.9	135	7	ABO68543	AbO68543 Pseudomon
162	34	60.7	155	7	ABO81847	AbO81847 Pseudomon	235	33	58.9	137	4	AAU42999	Aau42999 Propionib
163	34	60.7	159	7	ABO63703	AbO63703 Klebsiell	236	33	58.9	137	6	ABM39518	Abm39518 Propionib
164	34	60.7	193	9	ADY18608	Ady18608 PRO polyp	237	33	58.9	144	8	ADY23540	Ady23540 Plant ful
165	34	60.7	196	8	ADL06563	Adl06563 Human tum	238	33	58.9	156	4	ABG03602	Abg03602 Novel hum
166	34	60.7	202	7	ADB64982	Adb64982 Human pro	239	33	58.9	160	4	AAU44716	Aau44716 Propionib
167	34	60.7	213	4	AAB19897	Aab19897 Polyketid	240	33	58.9	160	6	ABM41235	Abm41235 Propionib
168	34	60.7	215	6	ADA34477	Ada34477 Acinetoba	241	33	58.9	164	7	ABO70547	AbO70547 Pseudomon
169	34	60.7	219	7	ABO67958	AbO67958 Pseudomon	242	33	58.9	179	4	AAU36304	Aau36304 Pseudomon
170	34	60.7	228	8	ADX79685	Adx79685 Plant ful	243	33	58.9	179	6	ABU38519	Abu38519 Protein e

244	33	58.9	196	8	ADO13837	Adol3837 Protein e	317	33	58.9	693	4	ABG28564	Abg28564 Novel hum
245	33	58.9	196	9	AEB37723	Aeb37723 L. pneumo	318	33	58.9	702	4	ABG27190	Abg27190 Novel hum
246	33	58.9	196	9	AEb41033	Aeb41033 L. pneumo	319	33	58.9	702	7	ADC33307	Adc33307 Human nov
247	33	58.9	199	8	ADM90889	Adm90889 Human pha	320	33	58.9	713	8	ADS21802	Ads21802 Bacterial
248	33	58.9	205	8	ADP30231	Adp30231 Human sec	321	33	58.9	754	8	ADS30847	Ads30847 Bacterial
249	33	58.9	205	8	ADP30232	Adp30232 Human sec	322	33	58.9	774	4	AAU36453	Aau36453 Pseudomon
250	33	58.9	210	7	ADF58756	Adf58756 Human pol	323	33	58.9	774	6	ABU38785	Abu38785 Protein e
251	33	58.9	218	6	ABU40896	Abu40896 Protein e	324	33	58.9	777	7	ABO74951	Abo74951 Pseudomon
252	33	58.9	218	8	ADN26233	Adn26233 Bacterial	325	33	58.9	790	7	ABM86794	Abm86794 Rice abio
253	33	58.9	222	7	ADF07151	Adf07151 Bacterial	326	33	58.9	802	2	AAR56550	Aar56550 Cold accl
254	33	58.9	232	8	ADN47634	Adn47634 Thermococ	327	33	58.9	826	6	ABU28255	Abu28255 Protein e
255	33	58.9	249	7	ADC01166	Adc01166 Enterohae	328	33	58.9	828	2	AAW04357	Aaw04357 E. coli p
256	33	58.9	252	4	ABG28561	Abg28561 Novel hum	329	33	58.9	850	2	AAR86955	Aar86955 E. coli p
257	33	58.9	256	8	ADX94555	Adx94555 Plant ful	330	33	58.9	858	6	ABU47698	Abu47698 Protein e
258	33	58.9	257	8	ADN26848	Adn26848 Bacterial	331	33	58.9	858	6	ABU15100	Abu15100 Protein e
259	33	58.9	257	8	ADN26596	Adn26596 Bacterial	332	33	58.9	858	6	ABU47261	Abu47261 Protein e
260	33	58.9	258	8	ADN62549	Adn62549 A. thalia	333	33	58.9	860	6	ABU32346	Abu32346 Protein e
261	33	58.9	261	2	AAW88462	Aaw88462 Bacillus	334	33	58.9	863	7	ABO64567	Abo64567 Klebsiell
262	33	58.9	261	8	ADS28510	Ads28510 Bacterial	335	33	58.9	901	9	AEB17427	Aeb17427 E. coli K
263	33	58.9	267	8	ADT58437	Adt58437 Plant pol	336	33	58.9	923	8	ABM84093	Abm84093 Human dia
264	33	58.9	272	4	AAG81225	Aag81225 Mycobacte	337	33	58.9	923	8	ADT04046	Adt04046 Human pro
265	33	58.9	272	5	AAU91152	Aau91152 Mycobacte	338	33	58.9	929	8	ADX96680	Adx96680 Plant ful
266	33	58.9	281	8	ADX76931	Adx76931 Plant ful	339	33	58.9	932	4	ABG22206	Abg22206 Novel hum
267	33	58.9	287	4	AAU44866	Aau44866 Propionib	340	33	58.9	972	8	ADR10276	Adr10276 Human pro
268	33	58.9	287	6	ABM41385	Abm41385 Propionib	341	33	58.9	1044	4	ABG30271	Abg30271 Novel hum
269	33	58.9	287	7	ADB80169	Adb80169 Mycobacte	342	33	58.9	1047	4	ABG24636	Abg24636 Novel hum
270	33	58.9	290	4	ABU53127	Abu53127 Human met	343	33	58.9	1089	7	ADM29411	Adm29411 Human nov
271	33	58.9	299	8	ADS23673	Ads23673 Bacterial	344	33	58.9	1118	2	AAW82395	Aaw82395 Human UBP
272	33	58.9	332	3	AAW53447	Aaw53447 Human col	345	33	58.9	1118	7	ADG10514	Adg10514 Human STA
273	33	58.9	333	6	ABU40148	Abu40148 Protein e	346	33	58.9	1118	7	ADM29413	Adm29413 Human nov
274	33	58.9	333	8	ADK14053	Adk14053 C. glutam	347	33	58.9	1118	8	ADH09574	Adh09574 Human hos
275	33	58.9	335	6	ADA48338	Ada48338 Rice prot	348	33	58.9	1118	9	ADY77616	Ady77616 Human UBP
276	33	58.9	342	8	ADK14055	Adk14055 C. glutam	349	33	58.9	1128	4	ABG11843	Abg11843 Novel hum
277	33	58.9	345	2	AA31651	Aay31651 Corynebac	350	33	58.9	1130	4	ABG22207	Abg22207 Novel hum
278	33	58.9	345	2	AA322645	Aay22645 3-PGDH pr	351	33	58.9	2260	8	ADK16027	Adk16027 Streptomy
279	33	58.9	348	7	ABO77897	Abo77897 Pseudomon	352	33	58.9	3362	8	ADK16028	Adk16028 Streptomy
280	33	58.9	363	4	AAG91433	Aag91433 C glutami	353	33	58.9	5836	8	ADK16042	Adk16042 Streptomy
281	33	58.9	366	6	ADB06642	Adb06642 Alloiococ	354	32.5	58.0	173	7	ADD19312	Add19312 Human sec
282	33	58.9	375	8	ADJ35062	Adj35062 Xylanase	355	32.5	58.0	205	7	ADD19275	Add19275 Human sec
283	33	58.9	396	4	AAU35375	Aau35375 Haemophil	356	32.5	58.0	405	4	AAU38919	Aau38919 C. tracho
284	33	58.9	396	6	ABU30159	Abu30159 Protein e	357	32.5	58.0	405	7	ADD42710	Add42710 Chlamydia
285	33	58.9	403	6	ABU44787	Abu44787 Protein e	358	32.5	58.0	566	3	AA387134	Aay87134 Human sec
286	33	58.9	405	8	ADX72956	Adx72956 Plant ful	359	32.5	58.0	566	4	AAE06111	Aae06111 Human gen
287	33	58.9	423	7	ABO69948	Abo69948 Pseudomon	360	32.5	58.0	566	5	ABG333933	Abg333933 Human sec
288	33	58.9	441	7	ADG10512	Adg10512 Human STA	361	32.5	58.0	566	6	ADA57360	Ada57360 Human sec
289	33	58.9	444	7	ABO78759	Abo78759 Pseudomon	362	32.5	58.0	566	6	ADA41241	Ada41241 Human sec
290	33	58.9	448	5	ABB92481	Abb92481 Herbicida	363	32.5	58.0	566	6	ABR48002	Abr48002 Human sec
291	33	58.9	451	8	ADK14054	Adk14054 C. glutam	364	32.5	58.0	567	2	AAW75770	Aaw75770 Human oxi
292	33	58.9	460	8	ADS26304	Ads26304 Bacterial	365	32.5	58.0	567	2	AA306480	Aay06480 Human tum
293	33	58.9	497	6	ADB06644	Adb06644 Alloiococ	366	32.5	58.0	567	3	AA399364	Aay99364 Human PRO
294	33	58.9	507	4	ABG28559	Abg28559 Novel hum	367	32.5	58.0	567	3	AA393687	Aay93687 Amino aci
295	33	58.9	513	6	ABU54545	Abu54545 Human NOV	368	32.5	58.0	567	4	AA366113	Aab66113 Protein o
296	33	58.9	517	4	AAU23711	Aau23711 Novel hum	369	32.5	58.0	567	4	AA372874	Aay72874 Human PRO
297	33	58.9	530	2	AA31649	Aay31649 Brevibact	370	32.5	58.0	567	4	AA350962	Aab50962 Human PRO
298	33	58.9	530	2	AA322646	Aay22646 Wild type	371	32.5	58.0	567	4	AAU12409	Aau12409 Human PRO
299	33	58.9	530	4	AAG91161	Aag91161 C glutami	372	32.5	58.0	567	5	ABB84904	Abb84904 Human PRO
300	33	58.9	530	8	ADK14058	Adk14058 C. glutam	373	32.5	58.0	567	5	ABB10104	Abb10104 Human imm
301	33	58.9	547	8	ADN17962	Adn17962 Bacterial	374	32.5	58.0	567	5	ABB95510	Abb95510 Human ang
302	33	58.9	553	9	ABM96964	Abm96964 M. xanthu	375	32.5	58.0	567	6	ABO17853	Abo17853 Novel hum
303	33	58.9	569	8	ADU07623	Adu07623 Amino aci	376	32.5	58.0	567	6	ADA56835	Ada56835 Human sec
304	33	58.9	592	8	ADX89308	Adx89308 Plant ful	377	32.5	58.0	567	6	ABU81107	Abu81107 Human PRO
305	33	58.9	602	8	ADY09874	Ady09874 Plant ful	378	32.5	58.0	567	6	ABU71428	Abu71428 Human neo
306	33	58.9	605	7	ABO77909	Abo77909 Pseudomon	379	32.5	58.0	567	6	ABU66807	Abu66807 Human PRO
307	33	58.9	608	6	ADB06646	Adb06646 Alloiococ	380	32.5	58.0	567	6	ADA40686	Ada40686 Human sec
308	33	58.9	608	7	ABO64692	Abo64692 Klebsiell	381	32.5	58.0	567	6	ABR47699	Abr47699 Human sec
309	33	58.9	622	8	ADN21005	Adn21005 Bacterial	382	32.5	58.0	567	6	ABU59888	Abu59888 Novel sec
310	33	58.9	634	7	ADG10510	Adg10510 Human STA	383	32.5	58.0	567	6	ABO25078	Abo25078 Human sec
311	33	58.9	634	7	ADG10508	Adg10508 Human STA	384	32.5	58.0	567	6	ABU67083	Abu67083 Human sec
312	33	58.9	650	6	ADB06648	Adb06648 Alloiococ	385	32.5	58.0	567	6	ADA45995	Ada45995 Novel hum
313	33	58.9	666	8	ADS43108	Ads43108 Bacterial	386	32.5	58.0	567	6	ADA76426	Ada76426 Human PRO
314	33	58.9	672	8	ADX76084	Adx76084 Plant ful	387	32.5	58.0	567	6	ADA19076	Ada19076 Human PRO
315	33	58.9	673	8	ADL05485	Adl05485 M. catarr	388	32.5	58.0	567	6	ADA61699	Ada61699 Homo sapi
316	33	58.9	682	6	ABU28397	Abu28397 Protein e	389	32.5	58.0	567	6	ADA61699	Ada61699 Homo sapi

390	32.5	58.0	567	6	ADB19484	Adb19484	Novel	hum	463	32.5	58.0	567	7	ADB36249	Adb36249	Human	PRO
391	32.5	58.0	567	6	ADB28025	Adb28025	Human	PRO	464	32.5	58.0	567	7	ADB46644	Adb46644	Novel	hum
392	32.5	58.0	567	6	ADA86504	Ada86504	Novel	hum	465	32.5	58.0	567	7	ADC17953	Adc17953	Human	PRO
393	32.5	58.0	567	6	ADB16068	Adb16068	Human	PRO	466	32.5	58.0	567	7	ADC50517	Adc50517	Novel	hum
394	32.5	58.0	567	6	ADA47854	Ada47854	Human	PRO	467	32.5	58.0	567	7	ADC72064	Adc72064	Novel	hum
395	32.5	58.0	567	6	ABO33606	Abo33606	Novel	hum	468	32.5	58.0	567	7	ADC60043	Adc60043	Novel	hum
396	32.5	58.0	567	6	ADA67649	Ada67649	Human	PRO	469	32.5	58.0	567	7	ADC53050	Adc53050	Novel	hum
397	32.5	58.0	567	6	ADB30656	Adb30656	Human	PRO	470	32.5	58.0	567	7	ADC57404	Adc57404	Novel	hum
398	32.5	58.0	567	6	ADA85952	Ada85952	Novel	hum	471	32.5	58.0	567	7	ADC60595	Adc60595	Novel	hum
399	32.5	58.0	567	6	ADA97164	Ada97164	Human	PRO	472	32.5	58.0	567	7	ADC51070	Adc51070	Novel	hum
400	32.5	58.0	567	6	ADA79468	Ada79468	Human	PRO	473	32.5	58.0	567	7	ADC65597	Adc65597	Human	PRO
401	32.5	58.0	567	6	ADA87607	Ada87607	Novel	hum	474	32.5	58.0	567	7	ADC54695	Adc54695	Novel	hum
402	32.5	58.0	567	6	ADB16809	Adb16809	Human	PRO	475	32.5	58.0	567	7	ADC53656	Adc53656	Novel	hum
403	32.5	58.0	567	6	ADA91901	Ada91901	Novel	hum	476	32.5	58.0	567	7	ADC59179	Adc59179	Novel	hum
404	32.5	58.0	567	6	ADB14964	Adb14964	Human	PRO	477	32.5	58.0	567	7	ADC56057	Adc56057	Novel	hum
405	32.5	58.0	567	6	ADB18925	Adb18925	Novel	hum	478	32.5	58.0	567	7	ADC58627	Adc58627	Novel	hum
406	32.5	58.0	567	6	ADA94140	Ada94140	Human	PRO	479	32.5	58.0	567	7	ADD03301	Add03301	Novel	hum
407	32.5	58.0	567	6	ADB20036	Adb20036	Novel	hum	480	32.5	58.0	567	7	ADC90293	Adc90293	Novel	hum
408	32.5	58.0	567	6	ADB13348	Adb13348	Human	PRO	481	32.5	58.0	567	7	ADC69712	Adc69712	Human	PRO
409	32.5	58.0	567	6	ABO43386	Abo43386	Novel	hum	482	32.5	58.0	567	7	ADC48601	Adc48601	Human	PRO
410	32.5	58.0	567	6	ADA74602	Ada74602	Human	PRO	483	32.5	58.0	567	7	ADD10130	Add10130	Human	PRO
411	32.5	58.0	567	6	ADB24835	Adb24835	Human	PRO	484	32.5	58.0	567	7	ADD04705	Add04705	Novel	hum
412	32.5	58.0	567	6	ADA82359	Ada82359	Human	PRO	485	32.5	58.0	567	7	ADC80661	Adc80661	Novel	hum
413	32.5	58.0	567	6	ADA75322	Ada75322	Human	PRO	486	32.5	58.0	567	7	ADD11168	Add11168	Human	PRO
414	32.5	58.0	567	6	ADA85400	Ada85400	Novel	hum	487	32.5	58.0	567	7	ADD10465	Add10465	Human	sec
415	32.5	58.0	567	6	ADA84848	Ada84848	Novel	hum	488	32.5	58.0	567	7	ADC48049	Adc48049	Human	PRO
416	32.5	58.0	567	6	ADB30104	Adb30104	Human	PRO	489	32.5	58.0	567	7	ADC80109	Adc80109	Novel	hum
417	32.5	58.0	567	6	ADA80632	Ada80632	Human	PRO	490	32.5	58.0	567	7	ADD11425	Add11425	Human	sec
418	32.5	58.0	567	6	ADA75874	Ada75874	Human	PRO	491	32.5	58.0	567	7	ADD09578	Add09578	Human	PRO
419	32.5	58.0	567	6	ADA47099	Ada47099	Human	PRO	492	32.5	58.0	567	7	ADD41291	Add41291	Novel	hum
420	32.5	58.0	567	6	ADB25395	Adb25395	Human	PRO	493	32.5	58.0	567	7	ADD52430	Add52430	Human	PRO
421	32.5	58.0	567	6	ADA93571	Ada93571	Human	PRO	494	32.5	58.0	567	7	ADD15372	Add15372	Novel	hum
422	32.5	58.0	567	6	ADB26921	Adb26921	Human	PRO	495	32.5	58.0	567	7	ADD70599	Add70599	Human	sec
423	32.5	58.0	567	6	ADB31208	Adb31208	Human	PRO	496	32.5	58.0	567	7	ADD39676	Add39676	Human	sec
424	32.5	58.0	567	6	ADA61136	Ada61136	Homo sapi		497	32.5	58.0	567	7	ADD53170	Add53170	Human	PRO
425	32.5	58.0	567	6	ADB24283	Adb24283	Human	PRO	498	32.5	58.0	567	7	ADD53722	Add53722	Novel	hum
426	32.5	58.0	567	6	ADA96612	Ada96612	Human	PRO	499	32.5	58.0	567	7	ADD70122	Add70122	Human	sec
427	32.5	58.0	567	6	ADA81184	Ada81184	Human	PRO	500	32.5	58.0	567	7	ADD37218	Add37218	Human	sec
428	32.5	58.0	567	6	ADA96060	Ada96060	Human	PRO	501	32.5	58.0	567	7	ADD38243	Add38243	Human	sec
429	32.5	58.0	567	6	ADB26369	Adb26369	Human	PRO	502	32.5	58.0	567	7	ADD39199	Add39199	Human	sec
430	32.5	58.0	567	6	ADB21854	Adb21854	Novel	hum	503	32.5	58.0	567	7	ADD51878	Add51878	Human	PRO
431	32.5	58.0	567	7	ADA77633	Ada77633	Human	PRO	504	32.5	58.0	567	7	ADD02677	Add02677	Human	PRO
432	32.5	58.0	567	7	ADB18373	Adb18373	Human	PRO	505	32.5	58.0	567	7	ADD02111	Add02111	Human	PRO
433	32.5	58.0	567	7	ADA87056	Ada87056	Novel	hum	506	32.5	58.0	567	7	ADD54293	Add54293	Novel	hum
434	32.5	58.0	567	7	ABO44459	Abo44459	Human	sec	507	32.5	58.0	567	7	ADD38722	Add38722	Human	sec
435	32.5	58.0	567	7	ADA88159	Ada88159	Novel	hum	508	32.5	58.0	567	7	ADD40153	Add40153	Human	sec
436	32.5	58.0	567	7	ADA46547	Ada46547	Novel	hum	509	32.5	58.0	567	7	ADE50374	Ade50374	Human	sec
437	32.5	58.0	567	7	ADB28577	Adb28577	Human	PRO	510	32.5	58.0	567	7	ADD92610	Add92610	Human	PRO
438	32.5	58.0	567	7	ADB29129	Adb29129	Human	PRO	511	32.5	58.0	567	7	ADD91506	Add91506	Human	PRO
439	32.5	58.0	567	7	ADA77081	Ada77081	Human	PRO	512	32.5	58.0	567	7	ADE04120	Ade04120	Human	PRO
440	32.5	58.0	567	7	ABO33483	Abo33483	Novel	hum	513	32.5	58.0	567	7	ADE19986	Ade19986	Human	sec
441	32.5	58.0	567	7	ADA88711	Ada88711	Novel	hum	514	32.5	58.0	567	7	ADE32417	Ade32417	Novel	hum
442	32.5	58.0	567	7	ADA97716	Ada97716	Human	PRO	515	32.5	58.0	567	7	ADE22349	Ade22349	Human	PRO
443	32.5	58.0	567	7	ADB27473	Adb27473	Human	PRO	516	32.5	58.0	567	7	ADD79573	Add79573	Human	PRO
444	32.5	58.0	567	7	ADB22406	Adb22406	Novel	hum	517	32.5	58.0	567	7	ADE42109	Ade42109	Human	PRO
445	32.5	58.0	567	7	ADA67097	Ada67097	Human	PRO	518	32.5	58.0	567	7	ADE17926	Adel17926	Human	PRO
446	32.5	58.0	567	7	ADB22958	Adb22958	Human	PRO	519	32.5	58.0	567	7	ADD92058	Add92058	Human	PRO
447	32.5	58.0	567	7	ADB23731	Adb23731	Human	PRO	520	32.5	58.0	567	7	ADE33521	Ade33521	Novel	hum
448	32.5	58.0	567	7	ADA92453	Ada92453	Novel	hum	521	32.5	58.0	567	7	ADE34073	Ade34073	Novel	hum
449	32.5	58.0	567	7	ADB15516	Adb15516	Human	PRO	522	32.5	58.0	567	7	ADD80125	Add80125	Human	PRO
450	32.5	58.0	567	7	ADB38768	Adb38768	Novel	hum	523	32.5	58.0	567	7	ADE49897	Ade49897	Human	sec
451	32.5	58.0	567	7	ADB38216	Adb38216	Novel	hum	524	32.5	58.0	567	7	ADD93162	Add93162	Human	PRO
452	32.5	58.0	567	7	ADB66688	Adb66688	Novel	hum	525	32.5	58.0	567	7	ADE19582	Ade19582	Human	PRO
453	32.5	58.0	567	7	ADB89768	Adb89768	Human	PRO	526	32.5	58.0	567	7	ADE21455	Ade21455	Human	sec
454	32.5	58.0	567	7	ADB90500	Adb90500	Human	PRO	527	32.5	58.0	567	7	ADE19030	Adel19030	Human	PRO
455	32.5	58.0	567	7	ADB39601	Adb39601	Novel	hum	528	32.5	58.0	567	7	ADE43226	Ade43226	Human	PRO
456	32.5	58.0	567	7	ADB47224	Adb47224	Novel	hum	529	32.5	58.0	567	7	ADD96015	Add96015	Human	PRO
457	32.5	58.0	567	7	ADB86831	Adb86831	Human	PRO	530	32.5	58.0	567	7	ADE22901	Ade22901	Human	PRO
458	32.5	58.0	567	7	ADB77436	Adb77436	Novel	hum	531	32.5	58.0	567	7	ADD79019	Add79019	Human	PRO
459	32.5	58.0	567	7	ADB34593	Adb34593	Human	PRO	532	32.5	58.0	567	7	ADE32969	Ade32969	Novel	hum
460	32.5	58.0	567	7	ADB35697	Adb35697	Human	PRO	533	32.5	58.0	567	7	ADE42661	Ade42661	Human	PRO
461	32.5	58.0	567	7	ADB34041	Adb34041	Human	PRO	534	32.5	58.0	567	7	ADD80677	Add80677	Human	PRO
462	32.5	58.0	567	7	ADB35145	Adb35145	Human	PRO	535	32.5	58.0	567	7	ADD89705	Add89705	Human	PRO

536	32.5	58.0	567	7	ADE40989	Ade40989 Human PRO	609	32.5	58.0	567	8	ADG62121	Adg62121 Novel hum
537	32.5	58.0	567	7	ADE04788	Ade04788 Human PRO	610	32.5	58.0	567	8	ADH02972	Adh02972 Human sec
538	32.5	58.0	567	7	ADE92917	Ade92917 Human PRO	611	32.5	58.0	567	8	ADG82322	Adg82322 Human PRO
539	32.5	58.0	567	7	ADF29880	Adf29880 Human sec	612	32.5	58.0	567	8	ADG57561	Adg57561 Novel hum
540	32.5	58.0	567	7	ADE94077	Ade94077 Immune di	613	32.5	58.0	567	8	ADG57009	Adg57009 Novel hum
541	32.5	58.0	567	7	ADF55773	Adf55773 Human sec	614	32.5	58.0	567	8	ADG55905	Adg55905 Novel hum
542	32.5	58.0	567	7	ADG21626	Adg21626 Novel hum	615	32.5	58.0	567	8	ADG58665	Adg58665 Novel hum
543	32.5	58.0	567	7	ADG23267	Adg23267 Novel hum	616	32.5	58.0	567	8	ADG71031	Adg71031 Novel hum
544	32.5	58.0	567	7	ADF97602	Adf97602 Human PRO	617	32.5	58.0	567	8	ADH03926	Adh03926 Human sec
545	32.5	58.0	567	7	ADG80666	Adg80666 Human PRO	618	32.5	58.0	567	8	ADH03449	Adh03449 Human sec
546	32.5	58.0	567	7	ADG80114	Adg80114 Human PRO	619	32.5	58.0	567	8	ADG58113	Adg58113 Novel hum
547	32.5	58.0	567	7	ADH55406	Adh55406 Novel hum	620	32.5	58.0	567	8	ADG53697	Adg53697 Novel hum
548	32.5	58.0	567	7	ADH55958	Adh55958 Novel hum	621	32.5	58.0	567	8	ADG71583	Adg71583 Novel hum
549	32.5	58.0	567	7	ADH99277	Adh99277 Human sec	622	32.5	58.0	567	8	ADG81770	Adg81770 Human PRO
550	32.5	58.0	567	7	ADI65126	Adi65126 Novel hum	623	32.5	58.0	567	8	ADH30732	Adh30732 Human PRO
551	32.5	58.0	567	7	ADI63625	Adi63625 Novel hum	624	32.5	58.0	567	8	ADH12099	Adh12099 Novel hum
552	32.5	58.0	567	7	ADH82039	Adh82039 Novel hum	625	32.5	58.0	567	8	ADG52521	Adg52521 Novel hum
553	32.5	58.0	567	7	ADH81487	Adh81487 Novel hum	626	32.5	58.0	567	8	ADG54249	Adg54249 Novel hum
554	32.5	58.0	567	7	ADM82656	Adm82656 Novel hum	627	32.5	58.0	567	8	ADG81218	Adg81218 Human PRO
555	32.5	58.0	567	7	ADN16055	Adn16055 Novel hum	628	32.5	58.0	567	8	ADG56457	Adg56457 Novel hum
556	32.5	58.0	567	7	ADN16684	Adn16684 Novel hum	629	32.5	58.0	567	8	ADH12723	Adh12723 Novel hum
557	32.5	58.0	567	7	ADN15503	Adn15503 Novel hum	630	32.5	58.0	567	8	ADG61569	Adg61569 Novel hum
558	32.5	58.0	567	7	ADN14951	Adn14951 Novel hum	631	32.5	58.0	567	8	ADH28656	Adh28656 Human PRO
559	32.5	58.0	567	7	ADI64177	Adi64177 Novel hum	632	32.5	58.0	567	8	ADG54801	Adg54801 Novel hum
560	32.5	58.0	567	8	ADC81213	Adc81213 Novel hum	633	32.5	58.0	567	8	ADH61404	Adh61404 Human sec
561	32.5	58.0	567	8	ADD76661	Add76661 Human PRO	634	32.5	58.0	567	8	ADG10008	Adg10008 Novel hum
562	32.5	58.0	567	8	ADD88025	Add88025 Human PRO	635	32.5	58.0	567	8	ADG115479	Adg115479 Novel hum
563	32.5	58.0	567	8	ADD86429	Add86429 Human PRO	636	32.5	58.0	567	8	ADG09356	Adg09356 Novel hum
564	32.5	58.0	567	8	ADE75877	Ade75877 Human PRO	637	32.5	58.0	567	8	ADI14811	Adi14811 Novel hum
565	32.5	58.0	567	8	ADE41426	Ade41426 Human sec	638	32.5	58.0	567	8	ADI18406	Adi18406 Novel hum
566	32.5	58.0	567	8	ADE23453	Ade23453 Human PRO	639	32.5	58.0	567	8	ADI37004	Adi37004 Novel hum
567	32.5	58.0	567	8	ADE24005	Ade24005 Human PRO	640	32.5	58.0	567	8	ADK52108	Adk52108 Human ato
568	32.5	58.0	567	8	ADE24648	Ade24648 Human PRO	641	32.5	58.0	567	8	ADJ63687	Adj63687 Novel hum
569	32.5	58.0	567	8	ADD87473	Add87473 Human PRO	642	32.5	58.0	567	8	ADJ77582	Adj77582 Human PRO
570	32.5	58.0	567	8	ADE89339	Ade89339 Human PRO	643	32.5	58.0	567	8	ADK82954	Adk82954 Human PRO
571	32.5	58.0	567	8	ADE18478	Ade18478 Human PRO	644	32.5	58.0	567	8	ADJ65704	Adj65704 Human PRO
572	32.5	58.0	567	8	ADE88787	Ade88787 Human PRO	645	32.5	58.0	567	8	ADM27840	Adm27840 Human PRO
573	32.5	58.0	567	8	ADE96457	Ade96457 Human sec	646	32.5	58.0	567	8	ADL82807	Adl82807 Human PRO
574	32.5	58.0	567	8	ADE94807	Ade94807 Human PRO	647	32.5	58.0	567	8	ADM42564	Adm42564 Human PRO
575	32.5	58.0	567	8	ADE91218	Ade91218 Human PRO	648	32.5	58.0	567	8	ADL91823	Adl91823 Human PRO
576	32.5	58.0	567	8	ADF25768	Adf25768 Human sec	649	32.5	58.0	567	8	ADL94603	Adl94603 Human sec
577	32.5	58.0	567	8	ADE95359	Ade95359 Human PRO	650	32.5	58.0	567	8	ADM28426	Adm28426 Human PRO
578	32.5	58.0	567	8	ADE93469	Ade93469 Human PRO	651	32.5	58.0	567	8	ADI95908	Adi95908 Human PRO
579	32.5	58.0	567	8	ADF24667	Adf24667 Human sec	652	32.5	58.0	567	8	ADI96460	Adi96460 Novel hum
580	32.5	58.0	567	8	ADF29403	Adf29403 Human PRO	653	32.5	58.0	567	8	ADP55860	Adp55860 Novel hum
581	32.5	58.0	567	8	ADF35050	Adf35050 Human PRO	654	32.5	58.0	567	8	ADS32412	Ads32412 Novel hum
582	32.5	58.0	567	8	ADE96934	Ade96934 Human sec	655	32.5	58.0	567	8	ADT03396	Adt03396 Human PRO
583	32.5	58.0	567	8	ADE92365	Ade92365 Novel hum	656	32.5	58.0	567	8	ADX57621	Adx57621 Rheumatoi
584	32.5	58.0	567	8	ADE90666	Ade90666 Human PRO	657	32.5	58.0	567	9	ADZ03447	Adz03447 Human sec
585	32.5	58.0	567	8	ADE91813	Ade91813 Novel hum	658	32.5	58.0	567	9	ADZ03447	Adz03447 Human sec
586	32.5	58.0	567	8	ADG23392	Adg23392 Human PRO	659	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
587	32.5	58.0	567	8	ADG22178	Adg22178 Novel hum	660	32.5	58.0	567	9	ADZ03447	Adz03447 Human sec
588	32.5	58.0	567	8	ADG20248	Adg20248 Human PRO	661	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
589	32.5	58.0	567	8	ADF98154	Adf98154 Human PRO	662	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
590	32.5	58.0	567	8	ADG24371	Adg24371 Novel hum	663	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
591	32.5	58.0	567	8	ADG24371	Adg24371 Novel hum	664	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
592	32.5	58.0	567	8	ADG03556	Adg03556 Human PRO	665	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
593	32.5	58.0	567	8	ADG03556	Adg03556 Human PRO	666	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
594	32.5	58.0	567	8	ADF99277	Adf99277 Human PRO	667	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
595	32.5	58.0	567	8	ADG16862	Adg16862 Human PRO	668	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
596	32.5	58.0	567	8	ADG05321	Adg05321 Human PRO	669	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
597	32.5	58.0	567	8	ADG19588	Adg19588 Human PRO	670	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
598	32.5	58.0	567	8	ADG13425	Adg13425 Human PRO	671	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
599	32.5	58.0	567	8	ADG08482	Adg08482 Novel hum	672	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
600	32.5	58.0	567	8	ADG15652	Adg15652 Human PRO	673	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
601	32.5	58.0	567	8	ADF97050	Adf97050 Human PRO	674	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
602	32.5	58.0	567	8	ADG06235	Adg06235 Human PRO	675	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
603	32.5	58.0	567	8	ADG23819	Adg23819 Novel hum	676	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
604	32.5	58.0	567	8	ADG04108	Adg04108 Human PRO	677	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
605	32.5	58.0	567	8	ADG25009	Adg25009 Novel hum	678	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
606	32.5	58.0	567	8	ADG07306	Adg07306 Novel hum	679	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
607	32.5	58.0	567	8	ADG07858	Adg07858 Novel hum	680	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
608	32.5	58.0	567	8	ADG55353	Adg55353 Novel hum	681	32.5	58.0	567	9	AEb14193	Aeb14193 Cancer ce
					ADG61017	Adg61017 Novel hum							

682	32	57.1	98	8	ADK69828	Adk69828	TIRAP pro	755	32	57.1	249	8	ADX93328	Adx93328	Plant ful
683	32	57.1	103	3	AAB42523	Aab42523	Human ORF	756	32	57.1	250	3	AAG18126	Aag18126	Arabidops
684	32	57.1	118	4	AAU47072	Aau47072	Propionib	757	32	57.1	251	2	AAR10668	Aar10668	Cationic
685	32	57.1	118	6	ABM43591	Abm43591	Propionib	758	32	57.1	251	2	AAR84666	Aar84666	Human hep
686	32	57.1	124	3	AAG13011	Aag13011	Arabidops	759	32	57.1	251	2	AAW88365	Aaw88365	Human pre
687	32	57.1	124	3	AAG50533	Aag50533	Arabidops	760	32	57.1	251	2	AAW88121	Aaw88121	Complete
688	32	57.1	129	3	AAG13010	Aag13010	Arabidops	761	32	57.1	251	3	AAW1878	Aay71878	Human pre
689	32	57.1	129	3	AAG50532	Aag50532	Arabidops	762	32	57.1	251	4	AAW1894	Aay71894	Human pre
690	32	57.1	136	3	AAG35509	Aag35509	Arabidops	763	32	57.1	251	9	ADX69318	Adx69318	Human hep
691	32	57.1	143	3	AAW95003	Aay95003	Human sec	764	32	57.1	259	3	AAG53308	Aag53308	Arabidops
692	32	57.1	143	4	AAB95555	Aab95555	Human pro	765	32	57.1	261	3	AAG06132	Aag06132	Arabidops
693	32	57.1	143	4	ABG02114	Abg02114	Novel hum	766	32	57.1	262	8	ADN25465	Adn25465	Bacterial
694	32	57.1	143	5	ABB98073	Abb98073	Human ary	767	32	57.1	262	8	ADY08479	Ady08479	Plant ful
695	32	57.1	143	6	ABU06886	Abu06886	Maize SSI	768	32	57.1	265	4	AAU36258	Aau36258	Pseudomon
696	32	57.1	146	6	ABU06685	Abu06685	Maize SSI	769	32	57.1	265	6	ABU38406	Abu38406	Protein e
697	32	57.1	147	3	AAG48729	Aag48729	Arabidops	770	32	57.1	266	8	ADY23109	Ady23109	Plant ful
698	32	57.1	147	3	AAG05246	Aag05246	Arabidops	771	32	57.1	272	7	ABO81531	Abo81531	Pseudomon
699	32	57.1	154	3	AAG07540	Aag07540	Arabidops	772	32	57.1	274	2	AAR84665	Aar84665	Human hep
700	32	57.1	154	5	ABG60122	Abg60122	Human DIT	773	32	57.1	277	6	ABM64806	Abm64806	Propionib
701	32	57.1	161	8	ADX90415	Adx90415	Plant ful	774	32	57.1	288	6	ABR42539	Abr42539	Clorobioc
702	32	57.1	165	8	ADT57682	Adt57682	Plant pol	775	32	57.1	289	8	ADY08001	Ady08001	Plant ful
703	32	57.1	167	7	ABO76469	AbO76469	Pseudomon	776	32	57.1	290	8	ADX73222	Adx73222	Plant ful
704	32	57.1	171	3	AAG35508	Aag35508	Arabidops	777	32	57.1	301	3	AAG13112	Aag13112	Arabidops
705	32	57.1	171	6	ABU43429	Abu43429	Protein e	778	32	57.1	309	3	AAG06131	Aag06131	Arabidops
706	32	57.1	172	6	ABU39620	Abu39620	Protein e	779	32	57.1	314	4	AAB70085	Aab70085	Human sec
707	32	57.1	182	3	AAG48728	Aag48728	Arabidops	780	32	57.1	314	8	ADY12461	Ady12461	Plant ful
708	32	57.1	182	3	AAG05245	Aag05245	Arabidops	781	32	57.1	319	3	AAG39923	Aag39923	Arabidops
709	32	57.1	182	5	ABB93582	Abb93582	Herbicida	782	32	57.1	321	3	AAG53307	Aag53307	Arabidops
710	32	57.1	185	4	AAU32997	Aau32997	Novel hum	783	32	57.1	324	3	AAG07122	Aag07122	Arabidops
711	32	57.1	191	3	AAG48727	Aag48727	Arabidops	784	32	57.1	326	8	ADX95574	Adx95574	Plant ful
712	32	57.1	191	3	AAG35507	Aag35507	Arabidops	785	32	57.1	328	8	ADX74803	Adx74803	Plant ful
713	32	57.1	192	3	AAG05244	Aag05244	Arabidops	786	32	57.1	329	3	AAG06130	Aag06130	Arabidops
714	32	57.1	197	1	AAP91931	Aap91931	Human hep	787	32	57.1	329	4	AAW78574	Aam78574	Human pro
715	32	57.1	198	3	AAB56131	Aab56131	Human sec	788	32	57.1	329	8	ADX97122	Adx97122	Plant ful
716	32	57.1	199	7	ADF17970	Adf17970	Human cat	789	32	57.1	330	8	ADX95575	Adx95575	Plant ful
717	32	57.1	199	7	ADG42089	Adg42089	Human cor	790	32	57.1	332	5	ABB08371	Abb08371	Y-FVW*7 a
718	32	57.1	200	9	ADV85561	Adv85561	Mutant ad	791	32	57.1	335	2	AAW15420	Aaw15420	Rat Spm-
719	32	57.1	207	7	ADC30968	Adc30968	Human nov	792	32	57.1	340	8	ADS24219	Ads24219	Bacterial
720	32	57.1	212	4	ABG23125	Abg23125	Novel hum	793	32	57.1	347	8	ADJ35032	Adj35032	Xylanase
721	32	57.1	216	8	ADX75990	Adx75990	Plant ful	794	32	57.1	349	8	ADS24608	Ads24608	Bacterial
722	32	57.1	217	6	ABU06500	Abu06500	Maize SSI	795	32	57.1	350	4	AAU61038	Aau61038	Propionib
723	32	57.1	221	2	AAR84663	Aar84663	Human hep	796	32	57.1	350	6	ABM57557	Abm57557	Propionib
724	32	57.1	221	2	AAW04875	Aaw04875	Heparin b	797	32	57.1	352	4	ABG08205	Abg08205	Novel hum
725	32	57.1	221	8	ADT57175	Adt57175	Plant pol	798	32	57.1	352	6	ABM68224	Abm68224	Photorhab
726	32	57.1	222	2	AAR41935	Aar41935	Recombina	799	32	57.1	357	3	ABG53306	Aag53306	Arabidops
727	32	57.1	222	2	AAW73210	Aaw73210	CAP37 pro	800	32	57.1	360	3	AAG07121	Aag07121	Arabidops
728	32	57.1	222	7	ADG42088	Adg42088	Human pol	801	32	57.1	360	7	ADE25082	Ade25082	Plant gro
729	32	57.1	225	2	AAW88362	Aaw88362	Human mat	802	32	57.1	361	7	ADJ70843	Adj70843	Human hea
730	32	57.1	225	2	AAW21551	Aay21551	Human hep	803	32	57.1	362	8	ADX87805	Adx87805	Plant ful
731	32	57.1	225	2	AAW88118	Aaw88118	Mature hu	804	32	57.1	363	8	ADT56787	Adt56787	Plant pol
732	32	57.1	225	3	AAW1881	Aay71881	Human hep	805	32	57.1	364	3	AAW44258	Aay44258	Reversibl
733	32	57.1	225	3	AAW1882	Aay71882	Human hep	806	32	57.1	366	2	AAW37453	Aay37453	Amino aci
734	32	57.1	225	3	AAW1883	Aay71883	Human hep	807	32	57.1	366	2	AAW35499	Aay35499	Chlamydia
735	32	57.1	225	3	AAW1876	Aay71876	Human mat	808	32	57.1	366	6	ABU27052	Abu27052	Protein e
736	32	57.1	225	4	AAW1891	Aay71891	Human mat	809	32	57.1	366	6	ABU27139	Abu27139	Protein e
737	32	57.1	225	7	ADE11587	Adel1587	Human mat	810	32	57.1	368	8	ADY24592	Ady24592	Plant ful
738	32	57.1	225	7	ADE11551	Adel1551	Human mat	811	32	57.1	368	8	ADX75138	Adx75138	Plant ful
739	32	57.1	225	8	ADK42019	Adk42019	Human hep	812	32	57.1	369	7	ADC68413	Adc68413	Lolium pe
740	32	57.1	225	9	ADV21099	Adv21099	Human azu	813	32	57.1	369	8	ADJ34800	Adj34800	Xylanase
741	32	57.1	225	9	ADX69310	Adx69310	Human hep	814	32	57.1	369	9	AEB03116	Aeb03116	Tannin bi
742	32	57.1	225	9	ADY82139	Ady82139	Human hep	815	32	57.1	371	5	ABB93226	Abb93226	Herbicida
743	32	57.1	227	5	AAU82726	Aau82726	Amino aci	816	32	57.1	372	8	ADN26117	Adn26117	Bacterial
744	32	57.1	228	7	ADE11545	Adel1545	Human pre	817	32	57.1	372	8	ADX67633	Adx67633	Plant ful
745	32	57.1	230	3	AAG50531	Aag50531	Arabidops	818	32	57.1	372	8	ADX68611	Adx68611	Plant ful
746	32	57.1	230	3	AAG13009	Aag13009	Arabidops	819	32	57.1	372	8	ADY05197	Ady05197	Plant ful
747	32	57.1	232	2	AAW88364	Aaw88364	Human pro	820	32	57.1	372	8	ADY09421	Ady09421	Plant ful
748	32	57.1	232	2	AAW21550	Aay21550	Human hep	821	32	57.1	372	8	ADM67627	Adm67627	Plant ful
749	32	57.1	232	2	AAW88120	Aaw88120	Human hep	822	32	57.1	372	9	ABM96610	Abm96610	M. xanthu
750	32	57.1	232	2	AAW88120	Aaw88120	Human hep	823	32	57.1	373	5	ABB92248	Abb92248	Herbicida
751	32	57.1	232	3	AAW1877	Aay71877	Human pre	824	32	57.1	374	8	ADY25279	Ady25279	Plant ful
752	32	57.1	236	7	ABO83170	AbO83170	Pseudomon	825	32	57.1	376	3	AAG07120	Aag07120	Arabidops
753	32	57.1	244	4	AAW1893	Aay71893	Human pre	826	32	57.1	376	8	ADT49303	Adt49303	Pectate l
754	32	57.1	245	8	ADN20995	Adn20995	Bacterial	827	32	57.1	378	5	ABB93127	Abb93127	Herbicida

828 32 57.1 380 4 AAM79558 Aam79558 Human pro
829 32 57.1 382 5 ABB09813 Abb09813 Amino aci
830 32 57.1 384 8 ADX87593 Adx87593 Plant ful
831 32 57.1 385 8 ADX93845 Adx93845 Plant ful
832 32 57.1 388 7 ABO78635 Abo78635 Pseudomon
833 32 57.1 390 7 ABO73621 Abo73621 Pseudomon
834 32 57.1 391 3 AAG13111 Aag13111 Arabidops
835 32 57.1 392 8 ADX72979 Adx72979 Plant ful
836 32 57.1 394 3 AAG13110 Aag13110 Arabidops
837 32 57.1 398 8 ADX67534 Adx67534 Plant ful
838 32 57.1 411 6 ABP70526 Abp70526 Histone d
839 32 57.1 412 7 ABO76339 Abo76339 Pseudomon
840 32 57.1 415 8 ADU06862 Adu06862 Cell adhe
841 32 57.1 417 5 ABB09808 Abb09808 Amino aci
842 32 57.1 417 8 ABM85095 Abm85095 Human dia
843 32 57.1 421 7 ADF94873 Adf94873 Human gen
844 32 57.1 421 7 ADF94874 Adf94874 Human gen
845 32 57.1 421 7 ADF94874 Adf94874 Human gen
846 32 57.1 421 8 ADH18875 Adh18875 Human cel
847 32 57.1 423 4 AAB92497 Aab92497 Human pro
848 32 57.1 428 6 ABJ25561 Abj25561 Aspergill
849 32 57.1 435 3 AAG39922 Aag39922 Arabidops
850 32 57.1 437 6 ABR41990 Abr41990 Eupenicil
851 32 57.1 437 9 ADU97708 Adu97708 E. terren
852 32 57.1 443 5 ABR40426 Abr40426 Human sec
853 32 57.1 443 5 ADI28010 Adi28010 ECMCAD pr
854 32 57.1 443 7 ADF94870 Adf94870 Human gen
855 32 57.1 445 8 ADM80779 Adm80779 Human CAD
856 32 57.1 445 8 ABM85093 Abm85093 Human dia
857 32 57.1 447 9 ADZ84977 Adz84977 FHOS prot
858 32 57.1 456 6 AAE38538 Aae38538 Partial b
859 32 57.1 458 7 ADF94875 Aaf94875 Human gen
860 32 57.1 459 4 AAB87356 Aab87356 Human gen
861 32 57.1 459 5 ABG65357 Abg65357 Human alb
862 32 57.1 459 7 ADF94869 Adf94869 Human gen
863 32 57.1 459 8 ADL78624 Adl78624 Albumin f
864 32 57.1 469 3 AAY44608 Aay44608 Maize MLO
865 32 57.1 471 3 AAB38627 Aab38627 Human sec
866 32 57.1 471 3 AAB38626 Aab38626 Gene 38 h
867 32 57.1 473 6 ABR58058 Abr58058 Human bes
868 32 57.1 473 6 ABR43181 Abr43181 Human REM
869 32 57.1 478 7 ADC68368 Adc68368 S. arundi
870 32 57.1 478 9 ABM90586 Abm90586 M. xanthu
871 32 57.1 478 9 AEB03071 Aeb03071 Fructan b
872 32 57.1 486 5 ADR41504 Adr41504 Human CD-
873 32 57.1 494 8 ADQ65690 Adq65690 Novel hum
874 32 57.1 509 7 ABO77666 Abo77666 Pseudomon
875 32 57.1 509 8 ADT87071 Adt87071 Yeast Str
876 32 57.1 520 8 ADQ65107 Adq65107 Novel hum
877 32 57.1 521 8 ADX91675 Adx91675 Plant ful
878 32 57.1 530 2 AAY31650 Aay31650 Brevibact
879 32 57.1 530 2 AAY22647 Aay22647 Mutant 3-
880 32 57.1 530 8 ADY13924 Ady13924 Plant ful
881 32 57.1 534 8 ADX94108 Adx94108 Plant ful
882 32 57.1 535 8 ADX73810 Adx73810 Plant ful
883 32 57.1 537 4 AAB95417 Aab95417 Human pro
884 32 57.1 543 3 AAG31951 Aag31951 Arabidops
885 32 57.1 544 5 ABB93224 Abb93224 Herbicida
886 32 57.1 544 7 ABO70318 Abo70318 Pseudomon
887 32 57.1 557 8 ADS24731 Ads24731 Bacterial
888 32 57.1 574 5 ABB93227 Abb93227 Herbicida
889 32 57.1 581 7 ADC68461 Adc68461 S. arundi
890 32 57.1 581 9 AEB03164 Aeb03164 Fructan b
891 32 57.1 587 5 ABB93801 Abb93801 Herbicida
892 32 57.1 587 8 ADN74203 Adn74203 Thale cre
893 32 57.1 592 9 AEA43084 Aea43084 M. smegma
894 32 57.1 592 9 AEA43085 Aea43085 M. smegma
895 32 57.1 609 3 AAG31950 Aag31950 Arabidops
896 32 57.1 610 7 ABO71553 Abo71553 Pseudomon
897 32 57.1 624 3 AAG32162 Aag32162 Arabidops
898 32 57.1 637 3 AAG39921 Aag39921 Arabidops
899 32 57.1 643 3 AAG32161 Aag32161 Arabidops
900 32 57.1 666 8 ADU07641 Adu07641 Amino aci

901 32 57.1 669 3 AAG31949 Aag31949 Arabidops
902 32 57.1 678 7 ADB65141 Adb65141 Human pro
903 32 57.1 688 4 ABG26120 Abg26120 Novel hum
904 32 57.1 695 4 AAB76535 Aab76535 Corynebac
905 32 57.1 699 9 ADW77609 Adw77609 Murine ac
906 32 57.1 700 3 AAY83636 Aay83636 Endoplasm
907 32 57.1 703 9 ADW77607 Adw77607 Human act
908 32 57.1 703 9 ADY70485 Ady70485 Human bet
909 32 57.1 703 9 ADZ28265 Adz28265 Human CRE
910 32 57.1 703 9 ADZ14931 Adz14931 Human CAM
911 32 57.1 736 6 ABU33299 Abu33299 Protein e
912 32 57.1 736 9 AEB36560 Aeb36560 L. pneumo
913 32 57.1 736 9 AEB39954 Aeb39954 L. pneumo
914 32 57.1 741 7 ADJ69381 Adj69381 Human hea
915 32 57.1 741 8 ADQ21134 Adq21134 Human sof
916 32 57.1 751 9 AEB48398 Aeb48398 Cassava s
917 32 57.1 754 8 ADN26732 Adn26732 Bacterial
918 32 57.1 754 8 ADN26587 Adn26587 Bacterial
919 32 57.1 758 7 ABO81228 Abo81228 Pseudomon
920 32 57.1 772 8 ADS26981 Ads26981 Bacterial
921 32 57.1 792 8 ADS26606 Ads26606 Bacterial
922 32 57.1 794 4 AAE03820 Aae03820 Human gen
923 32 57.1 794 5 ABG64543 Abg64543 Human alb
924 32 57.1 794 8 ADL77810 Adl77810 Albumin f
925 32 57.1 795 8 ADS27334 Ads27334 Bacterial
926 32 57.1 806 7 ADM04541 Adm04541 Human pro
927 32 57.1 811 5 ABP62957 Abp62957 Human pol
928 32 57.1 813 8 ADS23356 Ads23356 Bacterial
929 32 57.1 818 3 AAG32160 Aag32160 Arabidops
930 32 57.1 818 5 ABB92565 Abb92565 Herbicida
931 32 57.1 830 5 ABP62956 Abp62956 Human pol
932 32 57.1 869 8 ADR10147 Adr10147 Human pro
933 32 57.1 870 8 ADS28518 Ads28518 Bacterial
934 32 57.1 876 3 AAU28104 Aau28104 Human cyt
935 32 57.1 876 9 ADZ70761 Adz70761 PTPRF int
936 32 57.1 896 5 ABJ10550 Abj10550 Human NOV
937 32 57.1 896 8 ADH17468 Adh17468 Human NOV
938 32 57.1 914 8 ADH17466 Adh17466 Human NOV
939 32 57.1 915 2 AAY13350 Aay13350 Amino aci
940 32 57.1 915 3 AAY95340 Aay95340 Human PRO
941 32 57.1 915 3 ADC78354 Adc78354 Human PRO
942 32 57.1 915 4 AAB80218 Aab80218 Human PRO
943 32 57.1 915 4 AAU12318 Aau12318 Human PRO
944 32 57.1 915 4 AAB53077 Aab53077 Human ang
945 32 57.1 915 6 ABU71596 Abu71596 Human PRO
946 32 57.1 915 6 ABO17762 Abo17762 Novel hum
947 32 57.1 915 6 ABO17451 Abo17451 Human PRO
948 32 57.1 915 6 ABU81016 Abu81016 Human PRO
949 32 57.1 915 6 ABU71897 Abu71897 Human sec
950 32 57.1 915 6 ABU71897 Abu71897 Human sec
951 32 57.1 915 6 ABO01780 Abo01780 Novel hum
952 32 57.1 915 6 ABU66716 Abu66716 Human PRO
953 32 57.1 915 6 ABU54353 Abu54353 Human sec
954 32 57.1 915 6 ABO47368 Abo47368 Human sec
955 32 57.1 915 6 ABU59797 Abu59797 Novel sec
956 32 57.1 915 6 ABO24987 Abo24987 Human sec
957 32 57.1 915 6 ABU64505 Abu64505 Human sec
958 32 57.1 915 6 ABU67351 Abu67351 Human sec
959 32 57.1 915 6 ABO14871 Abo14871 Human sec
960 32 57.1 915 6 ABU66992 Abu66992 Human sec
961 32 57.1 915 6 ABU69628 Abu69628 Novel hum
962 32 57.1 915 6 ABO14810 Abo14810 Human sec
963 32 57.1 915 6 ADA45813 Ada45813 Novel hum
964 32 57.1 915 6 ADA76244 Ada76244 Human PRO
965 32 57.1 915 6 ADB29239 Adb29239 Human sec
966 32 57.1 915 6 ADA18894 Ada18894 Human PRO
967 32 57.1 915 6 ADA61517 Ada61517 Homo sapi
968 32 57.1 915 6 ADB19302 Adb19302 Novel hum
969 32 57.1 915 6 ADB27843 Adb27843 Human PRO
970 32 57.1 915 6 ADA86322 Ada86322 Novel hum
971 32 57.1 915 6 ADB15886 Adb15886 Human PRO
972 32 57.1 915 6 ADA47672 Ada47672 Human PRO
973 32 57.1 915 6 ADA18095 Ada18095 Human sec

974 32 57.1 915 6 ABO32762 Human sec
975 32 57.1 915 6 ADA67467 Human PRO
976 32 57.1 915 6 ADB30474 Human PRO
977 32 57.1 915 6 ADA85770 Novel hum
978 32 57.1 915 6 ADA96982 Human PRO
979 32 57.1 915 6 ADA79286 Human PRO
980 32 57.1 915 6 ADA87425 Novel hum
981 32 57.1 915 6 ADB16627 Human PRO
982 32 57.1 915 6 ABO34822 Human PRO
983 32 57.1 915 6 ADA16070 Human sec
984 32 57.1 915 6 ADA91719 Novel hum
985 32 57.1 915 6 ADB14782 Human PRO
986 32 57.1 915 6 ADB18743 Novel hum
987 32 57.1 915 6 ADA93958 Human PRO
988 32 57.1 915 6 ADB19854 Novel hum
989 32 57.1 915 6 ADB13166 Human PRO
990 32 57.1 915 6 ABO43295 Novel hum
991 32 57.1 915 6 ADA74420 Human PRO
992 32 57.1 915 6 ADA42215 Human sec
993 32 57.1 915 6 ADB24653 Human PRO
994 32 57.1 915 6 ADA82177 Human PRO
995 32 57.1 915 6 ADA75140 Human PRO
996 32 57.1 915 6 ADA85218 Novel hum
997 32 57.1 915 6 ADA84666 Novel hum
998 32 57.1 915 6 ABO17500 Human PRO
999 32 57.1 915 6 ADB29922 Human PRO
1000 32 57.1 915 6 ADA80450 Human PRO

ALIGNMENTS

RESULT 1
AAY22355
ID AAY22355 standard; peptide; 10 AA.

XX AAY22355;
AC
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 6.
XX

TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX Synthetic.
XX US95932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX

PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopaenia.
XX
PS Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration

XX Sequence 10 AA;

SQ Query Match 100.0%; Score 56; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 1 GRVRDQVAGW 10

RESULT 2
AAB16979
ID AAB16979 standard; peptide; 10 AA.

XX AAB16979;
AC
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:35.
XX

KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.

XX Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.

PS Claim 19; Page 207; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P'3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX

SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | | | | | | |
Db 1 GRVRDQVAGW 10

RESULT 3

ABB72865

ID ABB72865 standard; peptide; 10 AA.

XX

AC ABB72865;

XX

DT 05-APR-2002 (first entry)

XX

DE TPO mimetic peptide SEQ ID NO:35.

XX

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;

KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianaemic; anorectic; antiinfertility; haemostatic; dermatological;

KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

KW sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;

KW Fanconi's syndrome.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200183525-A2.

XX

PD 08-NOV-2001.

XX

PF 02-MAY-2001; 2001WO-US014310.

XX

PR 03-MAY-2000; 2000US-00563286.

XX

PA (AMGE-) AMGEN INC.

XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX

DR WPI; 2002-130313/17.

XX

PT Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX

PS Claim 39; Page 43; 176pp; English.

XX

CC The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteases of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

XX

SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | | | | | | |
Db 1 GRVRDQVAGW 10

RESULT 4

ADJ73016

ID ADJ73016 standard; peptide; 10 AA.

XX

AC ADJ73016;

XX

DT 06-MAY-2004 (first entry)

XX

DE TPO mimetic peptide sequence SeqID 470.

XX

KW mimetic; CDR mimetibody; gene therapy; transgenic; immune; cardiovascular; infectious; malignant; neurologic disease; anaemia;

KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective; TPO.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;

XX

DR WPI; 2003-804237/75.

XX

PT New CDR mimetibody comprising a portion of a heavy or light chain variable region comprising human framework or ligand binding region, useful for preparing a composition for treating e.g., immune, cardiovascular or neurologic disease.

XX

PS Disclosure; SEQ ID NO 470; 97pp; English.

XX

CC This invention relates to novel mammalian CDR mimetibodies, specific portions or variants thereof. Specifically, it refers to an antibody fragment where a protein has been inserted into, or replaces a portion of, one or more CDR regions, such that each CDR mimetibody comprises at least one portion of a heavy chain or light chain variable region, which itself comprises at least one human framework region and at least one ligand binding region (LBR). The present invention describes human mimetibodies, including modified immunoglobulins and cleavage products that can be useful in gene therapy and the generation of transgenic plants and animals. Furthermore, the CDR mimetibody is useful for preparing compositions for modulating, treating or reducing the symptoms of immune, cardiovascular, infectious, malignant and/ or neurologic diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 56; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 1 GRVRDQVAGW 10
|||||

RESULT 5
ADJ52651
ID ADJ52651 standard; peptide; 10 AA.
XX
AC ADJ52651;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID470.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 470; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 56; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 1 GRVRDQVAGW 10
|||||

RESULT 6
ADJ51612
ID ADJ51612 standard; peptide; 10 AA.
XX
AC ADJ51612;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID470.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 470; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|: |||||
Db 1 GRVRDQVAGW 10

RESULT 7
ADN20123
ID ADN20123 standard; protein; 780 AA.
XX
AC ADN20123;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #2776.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 2776; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 780 AA;

Query Match 73.2%; Score 41; DB 8; Length 780;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: |||||
Db 357 RLADQVAGW 365

RESULT 8
ADA36768
ID ADA36768 standard; protein; 787 AA.
XX
AC ADA36768;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #3929.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR N-PSDB; ADA32642.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 8055; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.

```
CC baumannii protein.
XX
SQ Sequence 787 AA;

Query Match      73.2%; Score 41; DB 6; Length 787;
Best Local Similarity 77.8%; Pred. NO. 58;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 RVRDQVAGW 10
      |:|||||
Db      387 RMADQVAGW 395

RESULT 9
AAM92258
ID AAM92258 standard; protein; 96 AA.
XX
AC AAM92258;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen SEQ ID NO: 1607.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.
XX
PN WC200155314-A2.
XX
PD 02-AUG-2001.
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
```


PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR N-PSDB; AAK88031.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Claim 11; SEQ ID NO 1607; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
SQ Sequence 96 AA;

Query Match 71.4%; Score 40; DB 4; Length 96;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db |:|:|:|
88 RIRDQLASW 96

RESULT 10
ABM91714
ID ABM91714 standard; protein; 353 AA.
XX
AC ABM91714;
XX
DT 02-JUN-2005 (first entry)
XX
DE M. xanthus protein sequence, seq id 10913.
KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX
OS Myxococcus xanthus.
XX
PN US6833447-B1.

XX 21-DEC-2004.
PD
XX
PF 10-JUL-2001; 2001US-00902540.
XX
PR 10-JUL-2000; 2000US-0217883P.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
DR WPI; 2005-028716/03.
XX
PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
PS Example 2; SEQ ID NO 10913; 25pp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 353 AA;

Query Match 71.4%; Score 40; DB 9; Length 353;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db |:|:|:|
21 GRIADTVSGW 30

RESULT 11
ADN21256
ID ADN21256 standard; protein; 363 AA.
XX
AC ADN21256;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #3909.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 3909; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 363 AA;

Query Match 71.4%; Score 40; DB 8; Length 363;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | |
Db 243 GEVPDQVVGW 252

RESULT 12
AAB54291
ID AAB54291 standard; protein; 95 AA.
XX
AC AAB54291;

XX 09-MAR-2001 (first entry)

XX Human pancreatic cancer antigen protein sequence SEQ ID NO:743.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
KW diagnosis; identification; cytostatic; neuroprotective; nootropic;
KW immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic; neural;
KW immune system; muscular; reproductive; gastrointestinal; pulmonary;
KW cardiovascular; renal; proliferative.

XX Homo sapiens.
OS
XX
XX WO200055320-A1.
PN
XX
XX 21-SEP-2000.
PD
XX

PF 08-MAR-2000; 2000WO-US005989.
XX
XX 12-MAR-1999; 99US-0124270P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2000-579444/54.
DR
DR N-PSDB; AAC99056.
XX

PT New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX

PS Claim 11; Page 1188; 1379pp; English.

XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX

SQ Sequence 95 AA;

Query Match 69.6%; Score 39; DB 3; Length 95;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | |
Db 46 GHRQRNIAGW 55

RESULT 13
AAB47456
ID AAB47456 standard; protein; 558 AA.
XX

AC AAB47456;

XX 13-DEC-2001 (first entry)

XX D-aminoacylase.

XX D-aminoacylase; N-acetyl-D-tryptophan; N-acetyl-D-phenylalanine; PCR;
KW N-acetyl-D-valine; N-acetyl-D-leucine; N-acetyl-D-methionine; primer;
KW probe; synthesis; detection; D-amino acid; substrate specific; amplify;
KW thermal stability; polymerase chain reaction.

XX Hypomyces mycophilus.
OS
XX
XX EP1120465-A1.
PN
XX
XX 01-AUG-2001.
PD
XX
XX 25-JAN-2001; 2001EP-00101739.
PF

XX 27-JAN-2000; 2000JP-00019080.
PR 22-MAY-2000; 2000JP-00150578.
XX
XX (DAIL) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Matsuyama A, Tokuyama S;
PI
XX WPI; 2001-551332/62.
DR N-PSDB; AAH43261.
XX
XX Novel D-aminoacylase-encoding gene derived from filamentous fungus
PT Hypomyces mycophilus, useful for producing D-tryptophan from N-acetyl-D-
PT tryptophan, useful as medicinal raw material.
XX
XX Claim 1; Page 17-20; 33pp; English.
PS
XX This sequence shows a D-aminoacylase polypeptide. D-aminoacylase has
CC physicochemical properties that include the action of the enzyme on N-
CC acetyl-D-amino acids to produce the corresponding D-amino acids, and
CC substrate specificity, where the enzyme acts on N-acetyl-D-tryptophan, N-
CC acetyl-D-phenylalanine, N-acetyl-D-valine, N-acetyl-D-leucine, and N-
CC acetyl-D-methionine, but not on N-acetyl-L-tryptophan, N-acetyl-L-
CC phenylalanine, N-acetyl-L-valine, N-acetyl-L-leucine, or N-acetyl-L-
CC methionine. Fragments of the D-aminoacylase cDNA are useful as primers or
CC as probes for synthesizing or detecting the full length cDNA. D-
CC aminoacylase is useful for producing D-amino acids. The enzyme has
CC substrate specificity, thermal stability and produces D-amino acid
CC efficiently by incubating the fungus derived D-aminoacylase with N-acetyl
CC -D-amino acid under proper conditions. The enzyme has high enzymatic
CC activity for N-acetyl-D-tryptophan and is excellent in industrial
CC applicability. The recombinant polypeptide of D-aminoacylase can be
CC manufactured at a low cost and in large quantities
XX
SQ Sequence 558 AA;

Query Match 69.6%; Score 39; DB 4; Length 558;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
Db 111 IREQIAGW 118

RESULT 14
ABO62198
ID ABO62198 standard; protein; 358 AA.
XX
AC ABO62198;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 8715.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.

DR N-PSDB; ACH95749.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 8715; 932pp; English.
PS
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 358 AA;

Query Match 67.9%; Score 38; DB 7; Length 358;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 272 GSLRDHVCGW 281

RESULT 15
AAW64365
ID AAW64365 standard; protein; 536 AA.
XX
AC AAW64365;
XX
DT 17-OCT-2003 (revised)
DT 09-NOV-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen XP14.
XX
KW Tuberculosis; infection; diagnosis; antigen; XP14.
XX
OS Mycobacterium tuberculosis; strain Erdman.
XX
PN WO9816645-A2.
XX
PD 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US018214.
XX
PR 11-OCT-1996; 96US-00729622.
PR 13-MAR-1997; 97US-00818111.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ;
XX
DR WPI; 1998-251292/22.
DR N-PSDB; AAV44415.
XX
PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
PT develop products for the detection of M. tuberculosis infection and
PT diagnosis of tuberculosis.
XX
PS Example 3; Page 173-175; 250pp; English.
XX
CC This polypeptide comprises Mycobacterium tuberculosis antigen XP14. Its
CC amino acid sequence was deduced from a DNA molecule (see AAV44415)
CC isolated from a M. tuberculosis strain Erdman genomic DNA expression
CC library using sera from patients having extrapulmonary tuberculosis. XP14
CC shows some sequence similarity to known sequences. The invention relates
CC to compositions and methods for diagnosing tuberculosis. It provides
CC polypeptides (see AAW64291-W64379) comprising an antigenic portion of a
CC soluble M. tuberculosis antigen, or an immunogenic portion of a M.
CC tuberculosis antigen, as well as DNA sequences encoding such

CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic kits for
CC detecting M. tuberculosis infection in a patient using these
CC polypeptides, antibodies or oligonucleotide probes and primers. (Updated
CC on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 335 GRIRDYVRSW 344

RESULT 16
AAW81732
ID AAW81732 standard; protein; 536 AA.
XX
AC AAW81732;
XX
DT 27-JAN-1999 (first entry)
XX
DE M. tuberculosis immunogenic polypeptide XP14 protein #1.
XX
KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW vaccine; pharmaceutical; infection; diagnosis.
XX
OS Mycobacterium tuberculosis.
XX WO9816646-A2.
XX
PD 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US018293.
XX
PR 11-OCT-1996; 96US-00730510.
PR 13-MAR-1997; 97US-00818112.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ;
XX
DR WPI; 1998-261042/23.
DR N-PSDB; AAV64524.
XX
PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
PT develop products for the detection of M. tuberculosis infection and for
PT diagnosis, treatment and prevention of tuberculosis.
XX
PS Example 3c; Page 164-165; 230pp; English.

CC This sequence represents an immunogenic portion of a soluble
CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for
CC inducing protective immunity against tuberculosis (TB). This sequence can
CC be formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis
XX
SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 335 GRIRDYVRSW 344

RESULT 17
AAV39019
ID AAY39019 standard; protein; 536 AA.
XX
AC AAY39019;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis recombinant antigen protein XP14.
XX
KW Antigen; diagnosis; detection; infection; antibody; immunisation;
KW vaccine; immunity.
XX
OS Mycobacterium tuberculosis.
XX WO9942118-A2.
PN
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US003265.
XX
PR 18-FEB-1998; 98US-00024753.
PR 05-MAY-1998; 98US-00072596.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR WPI; 1999-527416/44.
DR N-PSDB; AAZ19113.
XX
PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX
PS Example 3; Page 204-206; 323pp; English.

CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against M.
CC tuberculosis infection. The new detection methods are needed as current
CC vaccination strategies do not provide 100% immunity
XX
SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 335 GRIRDYVRSW 344

RESULT 18
AAV39162
ID AAY39162 standard; protein; 536 AA.
XX
AC AAY39162;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen XP14 amino acid sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942076-A2.
XX

PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US003268.
XX
PR 18-FEB-1998; 98US-00025197.
PR 05-MAY-1998; 98US-00072967.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR WPI; 1999-527409/44.
DR N-PSDB; AAZ19325.
XX
PT New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT tests and protective or therapeutic vaccines or compositions.
XX
PS Example 3; Page 159-161; 299pp; English.
XX
CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
CC polypeptides fragments, can be used in pharmaceutical compositions or
CC vaccines to generate a protective or therapeutic immune response to M.
CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural
CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
CC to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
CC the present invention
XX
SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:|||||
Db 335 GRIRDYRSW 344

RESULT 19
AAU54158
ID AAU54158 standard; protein; 103 AA.
XX
AC AAU54158;

XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #15054.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59563.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 15353; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 103 AA;

Query Match 66.1%; Score 37; DB 4; Length 103;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||||:|
Db 75 GRVRTAIRGW 84

RESULT 20
ABM50677
ID ABM50677 standard; protein; 103 AA.
XX
AC ABM50677;

XX 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #15353.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64492.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX

PS Example 1; SEQ ID NO 15353; 1481pp; English.

XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or

CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a

CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes

CC protein. The polynucleotides can also be used as probes or primers for

CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a polypeptide predicted to be encoded by an ORF (open

CC reading frame) contained within the P. acnes polynucleotides of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 103 AA;

Query Match 66.1%; Score 37; DB 6; Length 103;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||||| : |||
Db 75 GRVRTAIRGW 84

RESULT 21
ABO80351
ID ABO80351 standard; protein; 206 AA.
XX
AC ABO80351;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #12526.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.
DR N-PSDB; ABD13922.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 29097; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 206 AA;

Query Match 66.1%; Score 37; DB 7; Length 206;
Best Local Similarity 70.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | |||||
Db 5 GRSRTAVAGW 14

RESULT 22
ABO84669
ID ABO84669 standard; protein; 271 AA.
XX

AC ABO84669;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse cancer-associated protein MP20-011.1.
XX
KW Mouse; cancer-associated protein; cytostatic; cancer; leukaemia;
KW lymphoma; CAP.
XX
OS Mus musculus.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
DR WPI; 2004-652914/63.
DR N-PSDB; ABD32952.
XX

PT New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
PS disclosure; seqid 687; 310pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP protein sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 271 AA;

Query Match 66.1%; Score 37; DB 8; Length 271;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|::| |||
Db 202 GKXDSNAGW 211

RESULT 23
AAG92549
ID AAG92549 standard; protein; 297 AA.
XX
AC AAG92549;

XX 26-SEP-2001 (first entry)
DT
XX C glutamicum protein fragment SEQ ID NO: 6303.
DE
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
KW
XX Corynebacterium glutamicum.
OS
XX EP1108790-A2.
PN
XX 20-JUN-2001.
PD
XX 18-DEC-2000; 2000EP-00127688.
PF
XX 16-DEC-1999; 99JP-00377484.
PR
XX 07-APR-2000; 2000JP-00159162.
PR

PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH67768.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 6303; 246pp + Sequence Listing; English.
XX

CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 297 AA;

Query Match 66.1%; Score 37; DB 4; Length 297;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
||| |||
Db 74 RVRIQVTGW 82

RESULT 24
ADT57937
ID ADT57937 standard; protein; 423 AA.
XX
AC ADT57937;

XX 13-JAN-2005 (first entry)
DT
XX Plant polypeptide, SEQ ID 8014.
DE
XX

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
XX US2004216190-A1.
PN
XX 28-OCT-2004.
PD
XX 18-DEC-2003; 2003US-00739930.
PF
XX 28-APR-2003; 2003US-00424599.
PR
XX 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
DR WPI; 2004-757369/74.
XX

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.

PS Claim 2; SEQ ID NO 8014; 14pp; English.

XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX
SQ Sequence 423 AA;

Query Match 66.1%; Score 37; DB 8; Length 423;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 366 GMIRDQILSW 375

RESULT 25
AAG91159
ID AAG91159 standard; protein; 605 AA.

XX
AC AAG91159;

XX
DT 26-SEP-2001 (first entry)

XX
DE C glutamicum protein fragment SEQ ID NO: 4913.

XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.

XX
OS Corynebacterium glutamicum.

XX
PN EP1108790-A2.

XX
PD 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.
PF
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH66378.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.

PS Claim 17; SEQ ID NO 4913; 246pp + Sequence Listing; English.

XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX
SQ Sequence 605 AA;

Query Match 66.1%; Score 37; DB 4; Length 605;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 402 GRDTDQVAAW 411

RESULT 26
ADS29923
ID ADS29923 standard; protein; 746 AA.

XX
AC ADS29923;

XX
DT 02-DEC-2004 (first entry)

XX
DE Bacterial polypeptide #18956.

XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX
OS Bacteria.

XX
PN US2003233675-A1.

XX
PD 18-DEC-2003.

XX
PF 20-FEB-2003; 2003US-00369493.

XX
PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 18956; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 746 AA;

Query Match 66.1%; Score 37; DB 8; Length 746;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|: ||| ||
Db 344 RLADQVTCW 352

RESULT 27
AAU64755
ID AAU64755 standard; protein; 66 AA.
XX
AC AAU64755;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #25651.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX

PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59650.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 25950; 1069pp; English.
PS
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 66 AA;

Query Match 64.3%; Score 36; DB 4; Length 66;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
|| ||: ||
Db 9 GRPRDRRSGW 18

RESULT 28
ABM61274
ID ABM61274 standard; protein; 66 AA.
XX
AC ABM61274;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #25950.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX

PR 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
PA Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
PI
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64579.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 25950; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 66 AA;
Query Match 64.3%; Score 36; DB 6; Length 66;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
Db |||||:
9 GRPRDRRS GW 18
RESULT 29
AAB41784
ID AAB41784 standard; protein; 159 AA.
XX
AC AAB41784;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1548 polypeptide sequence SEQ ID NO:3096.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
XX WO200058473-A2.
PN
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
PI
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC75993.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 11; Page 2311; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 159 AA;
Query Match 64.3%; Score 36; DB 3; Length 159;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
Db |||||:
71 GRVVDVAGW 80
RESULT 30
ABM97186
ID ABM97186 standard; protein; 290 AA.
XX
AC ABM97186;
XX
DT 02-JUN-2005 (first entry)

XX DE M. xanthus protein sequence, seq id 16385.
XX KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX OS Myxococcus xanthus.
XX PN US6833447-B1.
XX PD 21-DEC-2004.
XX PF 10-JUL-2001; 2001US-00902540.
XX PR 10-JUL-2000; 2000US-0217883P.
XX PA (MONS) MONSANTO TECHNOLOGY LLC.
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX PT New substantially purified Myxococcus xanthus nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.
XX PS Example 2; SEQ ID NO 16385; 25pp; English.
XX CC The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX CC The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 9692-16825 represent
XX a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
SQ Sequence 290 AA;

Query Match 64.3%; Score 36; DB 9; Length 290;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|||:|
Db 191 RVRDEVRAW 199

RESULT 31
ABB84757
ID ABB84757 standard; protein; 336 AA.
XX
AC ABB84757;
XX
DT 16-MAY-2002 (first entry)
XX
DE DNA polymerase III holoenzyme delta subunit related protein SEQ ID:102.
XX
KW DNA polymerase III holoenzyme delta subunit; DNA polymerase III; DnaX;
KW hola; holB; antibacterial; tuberculostatic; antileprotic;
KW bacterial infection; DNA replication modulation.
XX Streptomyces coelicolor.
XX
PN WO200206532-A1.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022395.
XX
PR 14-JUL-2000; 2000US-0218246P.
PR 14-JUL-2000; 2000US-0218246P.

PR XX 28-MAR-2001; 2001US-00818780.
PA (REPL-) REPLIDYNE INC.
XX
PI Bullard JJ, Janjic N, Mchenry CS;
XX
DR WPI; 2002-164785/21.
DR N-PSDB; ABL88001.
XX
PT Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit from
PT bacteria, useful for screening agents that modulate the subunit activity
PT which is useful in the treatment of bacterial infections e.g. S. pyogenes
PT and S. aureus.
XX
PS Claim 19; Page 340-341; 500pp; English.
XX
CC The present invention describes nucleic acid sequences encoding a DNA
CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
CC antibacterial; tuberculostatic; antileprotic. Methods from the present
CC invention can be used for screening for bacterial DNA polymerase
CC holoenzyme delta subunit proteins and agents that modulate their
CC activity. The agents are useful in the treatment of bacterial infections
CC e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma, Yersinia,
CC Corynebacterium, Salmonella, Mycobacterium tuberculosis or M. leprae. The
CC invention provides a convenient means of identifying compounds which
CC modulate DNA replication in bacteria and therefore provide antibacterial
CC targets, and which are also useful for amplification of DNA. ABL87935 to
CC ABL88071 and ABB84724 to ABB84816 represent sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 336 AA;

Query Match 64.3%; Score 36; DB 5; Length 336;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|||:|
Db 281 RVRQMRGW 289

RESULT 32
ABB84815
ID ABB84815 standard; protein; 336 AA.
XX
AC ABB84815;
XX
DT 16-MAY-2002 (first entry)
XX
DE DNA polymerase III holoenzyme delta subunit related protein SEQ ID:229.
XX
KW DNA polymerase III holoenzyme delta subunit; DNA polymerase III; DnaX;
KW hola; holB; antibacterial; tuberculostatic; antileprotic;
KW bacterial infection; DNA replication modulation.
XX Streptomyces coelicolor.
XX
PN WO200206532-A1.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022395.
XX
PR 14-JUL-2000; 2000US-0218246P.
PR 28-MAR-2001; 2001US-00818780.
XX
PA (REPL-) REPLIDYNE INC.
XX
PI Bullard JJ, Janjic N, Mchenry CS;
XX
DR WPI; 2002-164785/21.
XX
PT Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit from

PT bacteria, useful for screening agents that modulate the subunit activity
PT which is useful in the treatment of bacterial infections e.g. S. pyogenes
PT and S. aureus.
XX
XX Claim 19; Fig 7F; 500pp; English.
XX
CC The present invention describes nucleic acid sequences encoding a DNA
CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
CC antibacterial; tuberculostatic; antileprotic. Methods from the present
CC invention can be used for screening for bacterial DNA polymerase
CC holoenzyme delta subunit proteins and agents that modulate their
CC activity. The agents are useful in the treatment of bacterial infections
CC e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma, Yersinia,
CC Corynebacterium, Salmonella, Mycobacterium tuberculosis or M. leprae. The
CC invention provides a convenient means of identifying compounds which
CC modulate DNA replication in bacteria and therefore provide antibacterial
CC targets, and which are also useful for amplification of DNA. ABL87935 to
CC ABL88071 and ABB84724 to ABB84816 represent sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 336 AA;

Query Match 64.3%; Score 36; DB 5; Length 336;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
||| |:
Db 281 RVRQMRGW 289

RESULT 33
ABG01473
ID ABG01473 standard; protein; 343 AA.
XX
AC ABG01473;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #1464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR
DR N-PSDB; AAS65660.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 31832; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 343 AA;

Query Match 64.3%; Score 36; DB 4; Length 343;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|:|
Db 89 GKLRPEAAGW 98

RESULT 34
ABU41540
ID ABU41540 standard; protein; 355 AA.
XX
AC ABU41540;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #27067.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Pseudomonas syringae.
XX
XX WO200277183-A2.
PN
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR
PR 06-SEP-2001; 2001US-00948993.
PR
PR 25-OCT-2001; 2001US-0342923P.
PR
PR 08-FEB-2002; 2002US-00072851.
PR
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI
XX WPI; 2003-029926/02.
DR
DR N-PSDB; ACA45410.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 59464; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 355 AA;

Query Match 64.3%; Score 36; DB 6; Length 355;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
| | : | : | | |
Db 168 GYVQRQIAGW 177

RESULT 35
ADH18901
ID ADH18901 standard; protein; 385 AA.
XX
AC ADH18901;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human cell adhesion and extracellular matrix CADECM-28 protein - SEQ 28.
XX
KW human; cell adhesion and extracellular matrix; CADECM; cytostatic;
KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
KW antiinflammatory; ophthalmological; antithyroid; antiarthritic;
KW antibacterial; virucide; protozoacide; antiparasitic; fungicide;
KW anorectic; cardiant; hypotensive; antiinfertility; hepatotropic;
KW autoimmune; inflammatory; AIDS; allergy; atopic dermatitis; arthritis;
KW thyroiditis; bacterial infection; viral; parasitic; protozoal; fungal;
KW metabolic; obesity; reproductive; infertility; neurological;
KW Parkinson's disease; Alzheimer's; stroke; epilepsy; cardiovascular;
KW myocardial infarction; hypertension; eye; cell proliferative; cancer;
KW atherosclerosis; hepatitis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
OS
XX
PN WO2003094843-A2.
XX
PD 20-NOV-2003.
XX
PF 06-MAY-2003; 2003WO-US014076.
XX
PR 10-MAY-2002; 2002US-0379840P.
PR 17-MAY-2002; 2002US-0381291P.
PR 24-MAY-2002; 2002US-0383183P.

PR 05-JUL-2002; 2002US-0394146P.
XX (INCY-) INCYTE CORP.
PA
XX Khare R, Elliott VS, Marquis JP, Ramkumar J, Chawla NK, Mason PM;
PI Hafalia AJA, Swarnakar A, Jin P, Becha SD, Kable AE, Tran UK;
PI Baughn MR, Burford N, Graul RC, Emerling BM, Sprague WW, Griffin JA;
PI Ison CH;
XX
DR WPI; 2004-011995/01.
DR N-PSDB; ADH18932.
XX
PT New human cell adhesion and extracellular matrix proteins and
PT polynucleotides, useful for diagnosing, preventing or treating diseases
PT or conditions associated with aberrant protein expression, e.g. cancer,
PT AIDS or stroke.
XX
PS Claim 1; SEQ ID NO 28; 308pp; English.
XX
CC The invention relates to a novel isolated human cell adhesion and
CC extracellular matrix (CADECM) polypeptide. The polypeptide of the
CC invention demonstrates cytostatic, antiarteriosclerotic, anti-HIV,
CC antiallergic, cerebroprotective, antiparkinsonian, anticonvulsant,
CC nootropic, neuroprotective, antiinflammatory, ophthalmological,
CC antithyroid, antiarthritic, antibacterial, virucide, protozoacide,
CC antiparasitic, fungicide, anorectic, cardiant, hypotensive,
CC antiinfertility and hepatotropic activities. The polypeptide may be
CC useful in diagnosing, preventing or treating diseases or conditions such
CC as autoimmune or inflammatory disorders including AIDS, allergies, atopic
CC dermatitis, arthritis or thyroiditis, infections particularly bacterial,
CC viral, parasitic, protozoal or fungal, metabolic disorders such as
CC obesity, reproductive disorders e.g. infertility, neurological disorders
CC including Parkinson's disease, Alzheimer's disease, stroke or epilepsy,
CC cardiovascular disorders e.g. myocardial infarction or hypertension, eye
CC disorders or cell proliferative diseases including cancer, the human
CC atherosclerosis and hepatitis. The current sequence is that of the human
CC CADECM protein of the invention.
XX
SQ Sequence 385 AA;

Query Match 64.3%; Score 36; DB 8; Length 385;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
| | | | | | | |
Db 348 GRVVDVADGW 357
RESULT 36
ABO81727
ID ABO81727 standard; protein; 387 AA.
XX
AC ABO81727;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #13902.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD15298.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 30473; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 387 AA;

Query Match 64.3%; Score 36; DB 7; Length 387;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | : | : | : |
Db 199 GYVQRQIAGW 208

RESULT 37
ABM15899
ID ABM15899 standard; protein; 415 AA.
XX
AC ABM15899;
XX
DT 26-SEP-2003 (first entry)
XX
DE Mycobacterium tuberculosis mycobacterial antigen protein SEQ ID NO:163.
XX
KW Mycobacterium tuberculosis; mycobacterial; antigen; infection; vaccine;
KW tuberculostatic; mycobacterial peptide; mycobacterial infection.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2003033530-A2.
XX
PD 24-APR-2003.
XX
PF 14-OCT-2002; 2002WO-GB004647.
XX
PR 12-OCT-2001; 2001GB-00024593.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI James B, Bacon J, March P;
XX
DR WPI; 2003-393501/37.
DR N-PSDB; ACF39393.
XX
PT New isolated mycobacterial peptide encoded by a gene that is induced or
PT up-regulated under high oxygen tension, useful for diagnosing, treating
PT or preventing a mycobacterial infection.

XX Claim 1; Page 269-270; 392pp; English.
PS
XX
CC The present invention describes an isolated mycobacterial peptide (I), or
CC its fragment, variant or derivative encoded by a gene whose expression is
CC induced or up-regulated during culture of a mycobacterium under
CC continuous culture conditions of a dissolved oxygen tension of at least
CC 30% air saturation measured at 37 plus degrees Celsius when compared with
CC a dissolved oxygen tension of up to 10% air saturation measured at 37
CC plus degrees Celsius. (I) has tuberculostatic activity and can be used in
CC vaccines. The mycobacterial peptide (I) or its fragment, variant or
CC derivative, inhibitor, antibody, attenuated mycobacterium, attenuated
CC microbial carrier, DNA sequence, DNA plasmid, RNA sequence, or RNA vector
CC from the present invention can be used for manufacturing a medicament for
CC treating or preventing a mycobacterial infection. The peptide or its
CC fragment, variant or derivative, the antibody, or a polynucleotide probe
CC comprising at least 8 nucleotides, where the probe binds to at least a
CC part of the gene, is useful for manufacturing a diagnostic reagent for
CC identifying a mycobacterial infection. The present sequence represents a
CC Mycobacterium tuberculosis mycobacterial antigen, which is used in the
CC exemplification of the present invention
XX
SQ Sequence 415 AA;

Query Match 64.3%; Score 36; DB 6; Length 415;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
| | : | : | : |
Db 174 VREQVIGW 181

RESULT 38
ABB89453
ID ABB89453 standard; protein; 423 AA.
XX
AC ABB89453;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 1829.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL89862.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 1829; 2081pp + Sequence Listing; English.
XX

CC The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB9040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 423 AA;

Query Match 64.3%; Score 36; DB 5; Length 423;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | | |
Db 179 GRVVDVADGW 188

RESULT 39
AAO30830
ID AAO30830 standard; protein; 436 AA.

XX AC AAO30830;

XX DT 22-SEP-2003 (first entry)

XX DE Human cell adhesion and extracellular matrix protein (CADECM)-20.

XX KW Human; cell adhesion and extracellular matrix protein; immune disorder; CADECM; cancer; gene therapy.

XX OS Homo sapiens.

XX PN WO2003047526-A2.

XX PD 12-JUN-2003.

XX PF 26-NOV-2002; 2002WO-US038437.

XX PR 30-NOV-2001; 2001US-0334343P.

PR 07-DEC-2001; 2001US-0340278P.

PR 04-JAN-2002; 2002US-0345069P.

PR 25-JAN-2002; 2002US-0351352P.

PR 14-FEB-2002; 2002US-0357168P.

PR 29-MAR-2002; 2002US-0369128P.

PR 05-APR-2002; 2002US-0370802P.

XX (INCY-) INCYTE GENOMICS INC.

PA Baughn MR, Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD; Delegeane AM, Elliott VS, Gandhi AR, Gietzen KJ, Gorvad AE; Griffin JA, Ho A, Jin P, Kable AE, Lal PG, Lee EA, Lee S, Lee SY; Marquis JP, Lehr-Mason PM, Ramkumar J, Richardson TW, Sprague WW; Swarnakar A, Tang TY, Tran B, Tran UK, Chawla NK, Warren BA, Xu Y; Yue H, Zheng W;

XX WPI; 2003-513695/48.

DR N-PSDB; AAL62032.

XX New human cell adhesion and extracellular matrix proteins (CADECM) polypeptide, useful for preparing a composition for treating a disease associated with decreased expression or overexpression of CADECM e.g.,

PT cancer.

XX Claim 1; Page 300-301; 374pp; English.

XX The invention relates to human cell adhesion and extracellular matrix proteins (CADECM) and nucleic acid molecules encoding such proteins. CADECM proteins are useful for preparing a composition for diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional CADECM e.g., immune disorders or cancer. The invention is also useful in gene therapy. The present sequence is human CADECM protein

XX SQ Sequence 436 AA;

Query Match 64.3%; Score 36; DB 7; Length 436;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | | |
Db 348 GRVVDVADGW 357

RESULT 40
ABB90892
ID ABB90892 standard; protein; 479 AA.

XX AC ABB90892;

XX DT 31-MAY-2002 (first entry)

XX DE Herbicidally active polypeptide SEQ ID NO 103.

XX KW Herbicidal; plant; agriculture; herbicide.

XX OS Arabidopsis thaliana.

XX PN WO200210210-A2.

XX PD 07-FEB-2002.

XX PF 28-AUG-2001; 2001WO-EP009892.

XX PR 28-AUG-2001; 2001WO-EP009892.

XX PA (FARB) BAYER AG.

XX PI Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from non-plant organisms.

XX Claim 5; SEQ ID NO 103; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides

XX SQ Sequence 479 AA;

Query Match 64.3%; Score 36; DB 5; Length 479;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RDQVAGW 10
|||: ||
Db 269 RDQIVGW 275

RESULT 41
ADT56368
ID ADT56368 standard; protein; 479 AA.
XX
AC ADT56368;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant polypeptide, SEQ ID 6445.
XX
KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
DR WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and genetics, and in particular for producing transgenic plants with improved biological characteristics.
PS Claim 2; SEQ ID NO 6445; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO: 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean, Arabidopsis, wheat and rape but the specification does not indicate which sequences is derived from which organism. Also included is a method of producing a plant having an improved property, comprising transforming a plant with a recombinant DNA construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide encoding a polypeptide associated with the property, and growing the transformed plant. The property is selected from improving plant cold tolerance, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for improving plant drought tolerance, for providing increased resistance to plant disease, for galactomannan production, for production of plant growth regulators, for improving plant heat tolerance, for improving plant tolerance to herbicides, for increasing the rate of homologous recombination in plants, for lignin production, for improving plant tolerance to extreme osmotic conditions, for improving plant tolerance to pathogens or pests, for yield improvement by modification of photosynthesis, for modifying seed oil yield and/or content, for modifying seed protein yield and/or content, for yield improvement by modification of carbohydrate, nitrogen or phosphorus use and/or uptake and for yield improvement by providing improved plant growth and development under at least one stress condition. The polynucleotide may also encode a plant transcription factor. The methods and compositions of the present invention are useful in the field of biochemistry and genetics, in particular for producing transgenic plants with improved biological characteristics such as increased yield, improved nitrogen

CC flow, increasing plant tolerance to cold or heat, improving plant tolerance to extreme osmotic and drought conditions, and improving plant tolerance to plant pests or pathogens. They can also be used in physical arrays of molecules, plant breeding markers, computer-based storage and analysis systems. The present sequence is one of the 5544 plant protein sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 479 AA;

Query Match 64.3%; Score 36; DB 8; Length 479;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RDQVAGW 10
|||: ||
Db 269 RDQIVGW 275

RESULT 42
ADV42198
ID ADV42198 standard; protein; 479 AA.
XX
AC ADV42198;
XX
DT 24-FEB-2005 (first entry)
XX
DE Glycosyltransferase UGT71C4 amino acid sequence.
XX
KW glycosyltransferase; protein engineering; enzyme engineering; enzyme.
XX
OS Unidentified.
XX
PN WO2004106508-A2.
XX
PD 09-DEC-2004.
XX
PF 24-MAY-2004; 2004WO-GB002237.
XX
PR 27-MAY-2003; 2003GB-00012042.
PR 28-JUN-2003; 2003GB-00015183.
XX
PA (UYYO-) UNIV YORK.
XX
PI Lim EK, Bowles D;
XX
DR WPI; 2005-021284/02.
DR N-PSDB; ADV42197.
XX
PT New reaction vessel comprising a genetically modified cell and a nutrient medium for supporting the growth of the cell that includes a substrate for glycosyltransferase, useful for producing recombinant proteins, e.g. proteases or amylase.
XX
PS Claim 6; Fig 3b; 72pp; English.
XX
CC The invention comprises a reaction vessel (e.g. a bioreactor) comprising a genetically modified cell, wherein the cell is modified by transfection or transformation with a nucleic acid that encodes a glycosyltransferase enzyme. The reaction vessel of the invention is useful for the production of molecules (e.g. recombinant proteins) on an industrial scale. The present amino acid sequence represents a glycosyltransferase enzyme of the invention.
XX
SQ Sequence 479 AA;

Query Match 64.3%; Score 36; DB 9; Length 479;
Best Local Similarity 71.4%; Pred. NO. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RDQVAGW 10

CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 3; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| |
Db 270 GRVVDADVGDW 279

RESULT 44
AAM93202
ID AAM93202 standard; protein; 517 AA.
XX
AC AAM93202;
XX
DT 06-NOV-2001 (first entry)
XX Human polypeptide, SEQ ID NO: 2591.
DE Human; full length cDNA; cDNA synthesis; oligo-capping.
XX Homo sapiens.
OS
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94110.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 2591; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 4; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| |
Db 270 GRVVDADVGDW 279

Db 269 RDQIVGW 275
|||: ||

RESULT 43
AAB41790
ID AAB41790 standard; protein; 517 AA.
XX
AC AAB41790;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1554 polypeptide sequence SEQ ID NO:3108.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
XX WO200058473-A2.
PN
PD 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shimkets RA, Leach M;
PI
XX WPI; 2000-602362/57.
DR N-PSDB; AAC75999.
DR
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 11; Page 2331-2332; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and

RESULT 45
AAM93206
ID AAM93206 standard; protein; 517 AA.
XX
AC AAM93206;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 2599.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94114.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2599; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 4; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 270 GRVVDVADGW 279

RESULT 46
ADL30558
ID ADL30558 standard; protein; 517 AA.
XX
AC ADL30558;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein encoded by a full length cDNA clone SeqID 2591.
XX

KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method.
XX
OS Homo sapiens.
XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR N-PSDB; ADL30557.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 1; SEQ ID NO 2591; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 8; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 270 GRVVDVADGW 279

RESULT 47
ADL30566
ID ADL30566 standard; protein; 517 AA.
XX
AC ADL30566;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein encoded by a full length cDNA clone SeqID 2599.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method.
XX
OS Homo sapiens.
XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
XX

PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR N-PSDB; ADL30565.
XX
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
PT
XX
PS Example 1; SEQ ID NO 2599; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 8; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | |
Db 270 GRVDAVDGW 279

RESULT 48
ADH18902
ID ADH18902 standard; protein; 527 AA.
XX
AC ADH18902;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human cell adhesion and extracellular matrix CADECM-29 protein - SEQ 29.
XX
KW human; cell adhesion and extracellular matrix; CADECM; cytostatic;
KW antiarteriosclerotic; anti-HIV; anti-allergic; cerebroprotective;
KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
KW anti-inflammatory; ophthalmological; antithyroid; antiarthritic;
KW antibacterial; virucide; protozoacide; antiparasitic; fungicide;
KW anorectic; cardiant; hypotensive; antiinfertility; hepatotropic;
KW autoimmune; inflammatory; AIDS; allergy; atopic dermatitis; arthritis;
KW thyroiditis; bacterial infection; viral; parasitic; protozoal; fungal;
KW metabolic; obesity; reproductive; infertility; neurological;
KW Parkinson's disease; Alzheimer's; stroke; epilepsy; cardiovascular;
KW myocardial infarction; hypertension; eye; cell proliferative; cancer;
KW atherosclerosis; hepatitis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX WO2003094843-A2.
PN
XX
PD 20-NOV-2003.
XX
XX 06-MAY-2003; 2003WO-US014076.
PF
XX 10-MAY-2002; 2002US-0379840P.
PR 17-MAY-2002; 2002US-0381291P.
PR 24-MAY-2002; 2002US-0383183P.
PR 05-JUL-2002; 2002US-0394146P.
XX

PA (INCY-) INCYTE CORP.
XX
XX Khare R, Elliott VS, Marquis JP, Ramkumar J, Chawla NK, Mason PM;
PI Hafalia AJA, Swarnakar A, Jin P, Becha SD, Kable AE, Tran UK;
PI Baughn MR, Burford N, Graul RC, Emerling BM, Sprague WM, Griffin JA;
PI Ison CH;
XX
DR WPI; 2004-011995/01.
DR N-PSDB; ADH18933.
XX
XX New human cell adhesion and extracellular matrix proteins and
PT polynucleotides, useful for diagnosing, preventing or treating diseases
PT or conditions associated with aberrant protein expression, e.g. cancer,
PT AIDS or stroke.
XX
PS Claim 1; SEQ ID NO 29; 308pp; English.
XX
CC The invention relates to a novel isolated human cell adhesion and
CC extracellular matrix (CADECM) polypeptide. The polypeptide of the
CC invention demonstrates cytostatic, antiarteriosclerotic, anti-HIV,
CC anti-allergic, cerebroprotective, antiparkinsonian, anticonvulsant,
CC nootropic, neuroprotective, antiinflammatory, ophthalmological,
CC antithyroid, antiarthritic, antiparasitic, anorectic, cardiant, hypotensive,
CC antiinfertility and hepatotropic activities. The polypeptide may be
CC useful in diagnosing, preventing or treating diseases or conditions such
CC as autoimmune or inflammatory disorders including AIDS, allergies, atopic
CC dermatitis, arthritis or thyroiditis, infections particularly bacterial,
CC viral, parasitic, protozoal or fungal, metabolic disorders such as
CC obesity, reproductive disorders e.g. infertility, neurological disorders
CC including Parkinson's disease, Alzheimer's disease, stroke or epilepsy,
CC cardiovascular disorders e.g. myocardial infarction or hypertension, eye
CC disorders or cell proliferative diseases including cancer.
CC atherosclerosis and hepatitis. The current sequence is that of the human
CC CADECM protein of the invention.
XX
SQ Sequence 527 AA;

Query Match 64.3%; Score 36; DB 8; Length 527;
Best Local Similarity 70.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | |
Db 348 GRVDAVDGW 357

RESULT 49
ADQ21291
ID ADQ21291 standard; protein; 560 AA.
XX
AC ADQ21291;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4111.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
XX WO2004048938-A2.
PN
XX 10-JUN-2004.
PD
XX 26-NOV-2003; 2003WO-US038193.
PF
XX 26-NOV-2002; 2002US-0429739P.
PR
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX

DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4111; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 560 AA;

Query Match 64.3%; Score 36; DB 8; Length 560;
Best Local Similarity 70.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | |
Db 347 GRVVDVADGW 356

RESULT 50
AAG81257
ID AAG81257 standard; protein; 578 AA.
XX
AC AAG81257;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human AFP protein sequence SEQ ID NO:32.
XX
KW Human; secreted protein; secretion; bacterial cell; fungal cell;
KW eukaryotic cell; fusion protein; maltose binding protein;
KW immunoglobulin constant region; polyhistidine tag.
XX
OS Homo sapiens.
XX
PN WO200129221-A2.
XX
PD 26-APR-2001.
XX
PF 20-OCT-2000; 2000WO-US029052.
XX
PR 20-OCT-1999; 99US-0160712P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC, Yee DP;
XX
DR WPI; 2001-300340/31.
DR N-PSDB; AAH52108.
XX
PT Isolated polypeptide for directing secretion of proteins of interest from
PT a host cell including, e.g. bacteria, includes contiguous amino acid
PT residues of polypeptide with specified amino acids.
XX
PS Claim 1; Page 121-122; 617pp; English.
XX
CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
CC to AAG81453. The secreted proteins can be used for directing the

CC secretion of proteins of interest from a host cell including bacteria,
CC fungal cells, and cultured higher eukaryotic cells. The present invention
CC also describes fusion proteins, where a secreted protein of the invention
CC is operably linked via a peptide bond or peptide linker to a second
CC protein selected from the group consisting of maltose binding protein, an
CC immunoglobulin constant region, a polyhistidine tag and a peptide given
CC in AAG81453
XX
SQ Sequence 578 AA;

Query Match 64.3%; Score 36; DB 4; Length 578;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| | | |
Db 348 GRVVDVADGW 357

Search completed: May 12, 2006, 10:34:53
Job time : 115.75 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-35
Perfect score: 56
Sequence: 1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: _pir1:*
2: _pir2:*
3: _pir3:*
4: _pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	41	73.2	780	1 S76487	probable copper-tr
2	39	69.6	320	2 E87684	ATP phosphoribosyl
3	39	69.6	360	2 AD1929	hypothetical prote
4	39	69.6	1216	2 T34101	hypothetical prote
5	38	67.9	300	2 B70725	probable linB prot
6	37	66.1	63	2 AB2809	conserved hypothet
7	37	66.1	181	2 C84233	hypothetical prote
8	37	66.1	218	1 S02828	albicidin resistan
9	37	66.1	421	2 C95880	probable protein l
10	37	66.1	627	2 AI2146	excinuclease ABC c
11	37	66.1	747	1 S36741	probable copper-tr
12	37	66.1	753	2 AH2543	cation-transportin
13	37	66.1	753	2 AE2009	cation-transportin
14	37	66.1	1752	2 A45407	collagen alpha 3(I
15	36	64.3	67	2 G82724	hypothetical prote
16	36	64.3	283	2 G82860	replication protei
17	36	64.3	338	2 C75353	ribosomal large su
18	36	64.3	356	2 E83416	hypothetical prote
19	36	64.3	415	2 F70686	hypothetical prote
20	36	64.3	479	2 G86207	hypothetical prote
21	36	64.3	723	2 I39066	N-methyl-D-asparta
22	36	64.3	951	2 G82965	conserved hypothet
23	36	64.3	1260	4 GNRL1	retrovirus-related
24	36	64.3	2517	2 S58380	probable RNA-direc
25	35	62.5	113	2 G82245	transcription regu
26	35	62.5	130	2 B69047	arsenate reductase
27	35	62.5	150	2 A41840	hypothetical prote
28	35	62.5	257	2 AD2083	phosphonate ABC tr
29	35	62.5	300	2 S76855	hypothetical prote

30	35	62.5	302	2	B87128	conserved hypothet
31	35	62.5	370	2	G83219	probable dihydrol
32	35	62.5	428	2	S38461	ubiquinol-cytochro
33	35	62.5	482	2	G84823	probable anthocyan
34	35	62.5	733	2	E95335	probable cation tr
35	35	62.5	891	2	H75507	hypothetical prote
36	35	62.5	1104	2	C72409	reverse gyrase - T
37	35	62.5	1273	2	E72611	probable ATP-depen
38	34	60.7	154	2	AC3434	calf thymus ribonu
39	34	60.7	186	2	C72782	hypothetical prote
40	34	60.7	193	2	D83620	hypothetical prote
41	34	60.7	213	2	G87457	guanylate kinase l
42	34	60.7	225	2	C84229	hypothetical prote
43	34	60.7	252	2	F87259	hypothetical prote
44	34	60.7	271	2	AI3628	peptidylprolyl iso
45	34	60.7	337	2	T16323	hypothetical prote
46	34	60.7	343	2	G71499	probable D-Ala-D-A
47	34	60.7	434	2	G95297	probable hydrolase
48	34	60.7	457	2	C65203	argininosuccinate
49	34	60.7	457	2	F86087	argininosuccinate
50	34	60.7	457	2	A98240	argininosuccinate
51	34	60.7	457	2	AI0477	probable argininos
52	34	60.7	458	2	AH0935	argininosuccinate
53	34	60.7	488	1	H64537	cytochrome-c oxida
54	34	60.7	488	2	G71969	cytochrome-c oxida
55	34	60.7	512	2	T30924	hypothetical prote
56	34	60.7	545	1	RRBPBM	RNA-directed RNA p
57	34	60.7	548	2	A39094	glucan endo-1,3-be
58	34	60.7	569	2	JC4177	Pah2 protein - yea
59	34	60.7	901	2	AG0996	Malt regulatory pr
60	34	60.7	1115	2	T29012	hypothetical prote
61	34	60.7	1166	2	F90247	reverse gyrase (to
62	34	60.7	1166	2	T29099	reverse gyrase - S
63	34	60.7	1177	2	G70399	reverse gyrase - A
64	34	60.7	1214	2	H75034	reverse gyrase (to
65	34	60.7	1221	2	T18550	reverse gyrase cha
66	34	60.7	1222	2	G72614	probable reverse g
67	34	60.7	1613	2	G64488	reverse gyrase (in
68	34	60.7	1624	2	C71129	probable reverse g
69	34	60.7	2022	2	T48818	glucan 1,4-alpha-g
70	33	58.9	68	2	AG3181	hypothetical prote
71	33	58.9	146	2	F72458	hypothetical prote
72	33	58.9	179	2	D83266	molybdopterin bios
73	33	58.9	190	2	C75055	transcription regu
74	33	58.9	225	2	B83083	conserved hypothet
75	33	58.9	240	2	F71236	hypothetical prote
76	33	58.9	244	2	D75505	hypothetical prote
77	33	58.9	249	2	F91095	type III secretion
78	33	58.9	249	2	B85941	hypothetical prote
79	33	58.9	261	2	E82748	tropinone reductas
80	33	58.9	272	2	A70955	hypothetical prote
81	33	58.9	275	2	C87634	conserved hypothet
82	33	58.9	287	2	H70923	probable hg transp
83	33	58.9	316	1	C69855	ABC transporter ho
84	33	58.9	357	2	E87584	transcription regu
85	33	58.9	360	2	C84243	asparagine synthet
86	33	58.9	384	2	AG3494	recF protein limpo
87	33	58.9	396	1	E64143	chloramphenicol re
88	33	58.9	396	2	E84196	cell division prot
89	33	58.9	432	2	B87161	competence damage
90	33	58.9	507	2	H86393	protein T24P13.8 l
91	33	58.9	541	2	T35143	probable monooxyge
92	33	58.9	547	2	G75364	DNA mismatch repai
93	33	58.9	630	2	T38023	probable transcrip
94	33	58.9	666	1	D69103	DNA helicase (EC 3
95	33	58.9	668	2	AG1989	hypothetical prote
96	33	58.9	679	2	S48939	hypothetical prote
97	33	58.9	707	2	A42322	ornithine decarbox
98	33	58.9	774	2	H83057	penicillin-binding
99	33	58.9	780	2	T09485	cold-induced prote
100	33	58.9	822	2	AB0238	heamin storage sys
101	33	58.9	822	2	T47007	hypothetical prote
102	33	58.9	858	1	ZPECPA	penicillin-binding

103 33 58.9 858 2 AE0999 penicillin-binding
104 33 58.9 858 2 D86004 peptidoglycan synt
105 33 58.9 858 2 F91158 peptidoglycan synt
106 33 58.9 901 1 RGE0MT regulatory protein
107 33 58.9 901 2 D91161 positive regulator
108 33 58.9 901 2 D86007 positive regulator
109 32.5 58.0 79 2 A87251 molybdopterin conv
110 32.5 58.0 176 2 G70599 hypothetical prote
111 32.5 58.0 398 2 A71482 hypothetical prote
112 32 57.1 73 2 F82172 hypothetical prote
113 32 57.1 75 2 S11156 psAS10 protein - c
114 32 57.1 138 2 T49392 related to pullula
115 32 57.1 158 2 H83405 hypothetical prote
116 32 57.1 176 2 T42179 hypothetical prote
117 32 57.1 182 2 T24757 hypothetical prote
118 32 57.1 186 2 AF2205 hypothetical prote
119 32 57.1 194 2 A75515 conserved hypothet
120 32 57.1 197 2 B71106 hypothetical prote
121 32 57.1 251 1 TRHUAZ azurocidin precurs
122 32 57.1 261 2 D96029 probable CoA-trans
123 32 57.1 265 2 C83419 NADH-dependent eno
124 32 57.1 286 2 H87506 hypothetical prote
125 32 57.1 301 2 H84192 hypothetical prote
126 32 57.1 306 2 T49541 hypothetical prote
127 32 57.1 322 2 T11576 type IIIa membrane
128 32 57.1 324 2 H72637 hypothetical prote
129 32 57.1 335 2 A49654 transcription fact
130 32 57.1 346 2 T11577 type IIIa membrane
131 32 57.1 360 2 T51394 reversibly glycosy
132 32 57.1 360 2 T33124 hypothetical prote
133 32 57.1 364 2 T04331 golgi associated p
134 32 57.1 364 2 T06507 reversibly glycosy
135 32 57.1 366 2 D72027 cell shape-determi
136 32 57.1 366 2 H86596 rod shape protein-
137 32 57.1 366 2 E81742 cell shape-determi
138 32 57.1 366 2 F71481 probable rod shape
139 32 57.1 366 2 AG0945 probable lipoprote
140 32 57.1 370 2 E71306 probable polynucle
141 32 57.1 370 2 C75268 carboxypeptidase G
142 32 57.1 370 2 B83480 probable oxidoredu
143 32 57.1 371 2 T05213 hypothetical prote
144 32 57.1 378 2 T02895 hypothetical prote
145 32 57.1 383 2 F83321 hypothetical prote
146 32 57.1 392 2 AD2360 hypothetical prote
147 32 57.1 402 2 AC2039 hypothetical prote
148 32 57.1 434 2 F84332 hypothetical prote
149 32 57.1 447 2 F69433 succinoglycan bios
150 32 57.1 465 2 AB3536 penicillin amidase
151 32 57.1 477 2 A87285 phosphate ABC tran
152 32 57.1 480 2 T50511 serine-type carbox
153 32 57.1 509 2 S74935 hypothetical prote
154 32 57.1 509 2 S30825 AFG1 protein - yea
155 32 57.1 539 2 B95313 cytochrome-c oxida
156 32 57.1 544 2 T05211 hypothetical prote
157 32 57.1 574 2 T05214 hypothetical prote
158 32 57.1 589 2 A82499 photoreceptor-inte
159 32 57.1 592 2 T48346 epithelial sodium
160 32 57.1 649 2 I64847 hypothetical prote
161 32 57.1 658 2 H95852 DNA helicase relat
162 32 57.1 669 2 F84433 probable cAMP-resp
163 32 57.1 690 2 C69208 hypothetical prote
164 32 57.1 699 2 T09069 catalase (EC 1.11.
165 32 57.1 709 2 T26874 hypothetical prote
166 32 57.1 727 2 F95422 ferrous iron trans
167 32 57.1 741 2 T46488 outer envelope mem
168 32 57.1 749 2 E75422 hypothetical prote
169 32 57.1 818 2 T12975 glycerol-3-phospha
170 32 57.1 840 2 F84213 sensor histidine k
171 32 57.1 870 2 B82732 hypothetical prote
172 32 57.1 892 2 A87447 Ca2+-transporting
173 32 57.1 935 2 E96806 Ca2+-transporting
174 32 57.1 994 2 A48849
175 32 57.1 1001 1 PWRBFC

176 32 57.1 1020 2 D83679
177 32 57.1 1036 2 A29832
178 32 57.1 1067 2 T18196
179 32 57.1 1146 2 B70376
180 32 57.1 1220 2 AD0125
181 32 57.1 1243 2 JC5615
182 32 57.1 1466 2 T30566
183 32 57.1 1621 2 S62356
184 32 57.1 1702 2 E86318
185 32 57.1 2108 2 H70819
186 32 57.1 3069 2 H70656
187 32 57.1 3076 2 A87058
188 32 57.1 3104 2 S20473
189 32 57.1 3341 1 A42996
190 31.5 56.2 246 2 C81946
191 31.5 56.2 250 2 C83397
192 31.5 56.2 327 2 T35322
193 31.5 56.2 747 2 D95862
194 31 55.4 57 2 C97937
195 31 55.4 114 2 T30727
196 31 55.4 140 2 S20914
197 31 55.4 140 2 S41577
198 31 55.4 140 2 S41574
199 31 55.4 141 2 S20915
200 31 55.4 148 1 FXDV
201 31 55.4 151 2 F69931
202 31 55.4 157 2 B70892
203 31 55.4 168 2 C70824
204 31 55.4 188 1 F64986
205 31 55.4 188 2 E85856
206 31 55.4 188 2 C91012
207 31 55.4 190 2 AF0784
208 31 55.4 192 2 T28569
209 31 55.4 194 2 A36851
210 31 55.4 194 2 A72167
211 31 55.4 200 2 E85047
212 31 55.4 219 2 G84388
213 31 55.4 221 2 D84516
214 31 55.4 226 2 F90469
215 31 55.4 228 2 H95868
216 31 55.4 229 2 H75594
217 31 55.4 240 2 D75203
218 31 55.4 241 2 B84992
219 31 55.4 247 2 A96001
220 31 55.4 253 2 G86605
221 31 55.4 253 2 H72017
222 31 55.4 255 2 AB3497
223 31 55.4 256 2 G70788
224 31 55.4 257 2 B75099
225 31 55.4 280 2 H70942
226 31 55.4 280 2 T03543
227 31 55.4 285 2 S65838
228 31 55.4 285 2 C84170
229 31 55.4 292 2 A75409
230 31 55.4 299 2 E87049
231 31 55.4 301 1 VHBPLL
232 31 55.4 301 2 T03320
233 31 55.4 302 2 S55644
234 31 55.4 310 2 C83267
235 31 55.4 318 2 G84143
236 31 55.4 330 2 E98119
237 31 55.4 343 2 D75260
238 31 55.4 345 2 A97113
239 31 55.4 350 2 D83102
240 31 55.4 363 2 AC0244
241 31 55.4 365 2 A34895
242 31 55.4 372 1 G70817
243 31 55.4 376 1 DCCEOC
244 31 55.4 386 2 G70779
245 31 55.4 399 2 T40877
246 31 55.4 400 2 T46526
247 31 55.4 410 2 B55523
248 31 55.4 412 2 C86177

hypothetical prote
HPI layer surface
pol protein - silk
reverse gyrase - A
exodeoxyribonuclea
membrane-associate
ATP-binding casset
TRP-185 protein -
protein F15H18.4 [
probable polyketid
fatty-acid synthas
fatty acid synthas
fatty-acid synthas
genome polyprotein
hypothetical prote
pyrroloquinoline q
hypothetical prote
probable sensor hi
transposase, uncha
hypothetical prote
lysozyme (EC 3.2.1
lysozyme (EC 3.2.1
lysozyme (EC 3.2.1
lysozyme (EC 3.2.1
flavodoxin - Desul
hypothetical prote
hypothetical prote
probable transcrip
probable lipoprote
probable lipoprote
probable lipoprote
probable lipoprote
A27L protein - var
A28L protein - var
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable osmotical
hypothetical prote
hypothetical prote
peptidylprolyl iso
conserved hypothet
snGlycerol-3-P acy
snGlycerol-3-p acy
methyltransferase
hypothetical prote
hypothetical prote
hypothetical prote
probable cobalt tr
dihydropteroate sy
ribosomal protein
carboxymethylenebu
hypothetical prote
major capsid prote
major capsid prote
hypothetical prote
hypothetical prote
hypothetical prote
transposase, uncha
probable carbohydr
probable Fe-S-clus
xenobiotic reducta
conserved hypothet
5S RNA-binding pro
probable ABC trans
orotidine-5'-phosp
hypothetical prote
hypothetical prote
probable hydroxyla
aromatic dioxygena
hypothetical prote

249 31 55.4 415 2 D96759 probable serine ca
250 31 55.4 416 2 B64132 tRNA adenyllyltrans
251 31 55.4 431 2 AG1884 glycolate oxidase
252 31 55.4 432 2 T36482 probable aminopept
253 31 55.4 433 2 B95405 probable nitritotr
254 31 55.4 435 2 E84618 probable serine ca
255 31 55.4 440 2 B87311 conserved hypothet
256 31 55.4 441 2 A96759 protein serine car
257 31 55.4 443 2 S66040 serine-type D-Ala-
258 31 55.4 443 2 JC4088 chorionic gonadotr
259 31 55.4 443 2 T45574 anthranilate N-hyd
260 31 55.4 448 2 A55519 mcrA protein - Str
261 31 55.4 468 2 H70521 probable papA2 pro
262 31 55.4 477 1 DYHUD5 dopamine receptor
263 31 55.4 483 1 WZBE20 gene 20 protein -
264 31 55.4 492 2 B86253 hypothetical prote
265 31 55.4 496 2 B84974 N utilization subs
266 31 55.4 527 2 A70975 hypothetical prote
267 31 55.4 532 2 A35149 ipaH protein - Shi
268 31 55.4 543 2 T32109 hypothetical prote
269 31 55.4 554 2 E83071 probable ATP-bind
270 31 55.4 555 2 G86138 ABC transporter in
271 31 55.4 555 2 E91297 ABC transporter in
272 31 55.4 555 2 F65254 ABC transporter in
273 31 55.4 555 2 AE1074 conserved hypothet
274 31 55.4 556 2 G64169 hypothetical prote
275 31 55.4 561 2 AG2336 potassium-dependen
276 31 55.4 564 2 E87631 acyl-CoA dehydroge
277 31 55.4 565 2 T38941 RNA binding protei
278 31 55.4 574 2 B35149 ipaH protein - Shi
279 31 55.4 580 2 AB0994 gamma-glutamyltran
280 31 55.4 585 2 F84346 archaeosine tRNA-r
281 31 55.4 587 2 AD3519 2,3-dihydroxybenzo
282 31 55.4 594 2 T31744 hypothetical prote
283 31 55.4 601 2 T36411 probable ABC-type
284 31 55.4 607 1 HMNZPD hemagglutinin - ph
285 31 55.4 607 1 JQ1535 hemagglutinin - ph
286 31 55.4 616 2 AB0824 chaperone protein
287 31 55.4 648 2 C83023 probable oxidoredu
288 31 55.4 660 2 D87331 TonB-dependent rec
289 31 55.4 690 2 T47478 hypothetical prote
290 31 55.4 698 2 T51915 hypothetical prote
291 31 55.4 720 2 S71130 catalase (EC 1.11.
292 31 55.4 722 2 D89701 protein F28H6.4 [i
293 31 55.4 733 2 G84668 hypothetical prote
294 31 55.4 735 2 F85055 hypothetical prote
295 31 55.4 743 2 T32794 hypothetical prote
296 31 55.4 754 2 S75113 catalase (EC 1.11.
297 31 55.4 771 2 A83348 probable aldehyde
298 31 55.4 781 2 F82584 catalase/peroxidas
299 31 55.4 809 2 S55344 outer envelope mem
300 31 55.4 836 2 S49940 cell division cont
301 31 55.4 885 1 S78246 endopeptidase Clp
302 31 55.4 903 2 AI0015 maltose regulon po
303 31 55.4 907 2 S54353 inter-alpha-trypsi
304 31 55.4 909 2 T06635 hypothetical prote
305 31 55.4 911 2 A39967 inter-alpha-trypsi
306 31 55.4 916 2 T35680 probable regulator
307 31 55.4 921 2 D82513 malt regulatory pr
308 31 55.4 944 2 C70839 probable mmpL3 pro
309 31 55.4 947 2 H85088 hypothetical prote
310 31 55.4 983 2 G84524 probable disease r
311 31 55.4 993 2 F97717 hypothetical prote
312 31 55.4 1000 2 T21522 hypothetical prote
313 31 55.4 1038 1 WZBE21 gene 21 protein -
314 31 55.4 1042 2 S43904 hyaluronidase - Cl
315 31 55.4 1080 2 E87586 metal ion efflux R
316 31 55.4 1120 2 H88449 protein F54D8.1 [i
317 31 55.4 1140 2 T09486 hypothetical prote
318 31 55.4 1178 2 AC3394 DNA-directed DNA p
319 31 55.4 1329 2 A64828 cell division prot
320 31 55.4 1342 2 E85614 cell division prot
321 31 55.4 1342 2 G90750 cell division prot

322 31 55.4 1343 2 AF0611 cell division prot
323 31 55.4 1516 2 F83085 conserved hypothet
324 31 55.4 1564 2 S55517 probable transport
325 31 55.4 1582 2 E70876 probable polyketid
326 31 55.4 1619 2 T30541 ABC1 transport pro
327 31 55.4 1647 2 T49412 hypothetical prote
328 31 55.4 1974 2 T30010 hypothetical prote
329 31 55.4 4063 2 T42993 probable spectrin
330 31 55.4 4101 2 T23630 hypothetical prote
331 31 55.4 15281 2 S41309 cyclosporin synthe
332 30.5 54.5 333 1 A38094 D-lactate dehydrog
333 30.5 54.5 337 1 S29296 probable substrate
334 30.5 54.5 542 2 T35938 55K protein precur
335 30.5 54.5 542 2 S07386 hypothetical prote
336 30.5 54.5 583 2 H69165 hypothetical prote
337 30.5 54.5 604 2 T36966 hypothetical prote
338 30.5 54.5 611 2 F82442 probable peptide A
339 30.5 54.5 639 2 H70661 probable dnaG prot
340 30.5 54.5 1171 2 F83110 exodeoxyribonuclea
341 30 53.6 46 2 F81142 hypothetical prote
342 30 53.6 50 2 S40969 TyB protein - yeas
343 30 53.6 80 2 A60699 hypothetical prote
344 30 53.6 80 2 D69745 hypothetical prote
345 30 53.6 101 2 D42074 hypothetical prote
346 30 53.6 104 2 B84037 thiorodoxin trxA l
347 30 53.6 105 2 A97519 hypothetical prote
348 30 53.6 120 2 D83166 hypothetical prote
349 30 53.6 127 2 S40034 killer toxin KP4 p
350 30 53.6 130 2 AC2735 hypothetical prote
351 30 53.6 146 2 H84358 hypothetical prote
352 30 53.6 150 2 A83217 conserved hypothet
353 30 53.6 158 2 F95415 hypothetical prote
354 30 53.6 160 2 A82020 probable periplasm
355 30 53.6 160 2 D81247 thiorodoxin-relate
356 30 53.6 162 2 D97821 hypothetical prote
357 30 53.6 184 2 G87321 hypothetical prote
358 30 53.6 189 2 A85489 hypothetical prote
359 30 53.6 189 2 A99638 hypothetical prote
360 30 53.6 194 2 AI0155 probable lipoprote
361 30 53.6 209 2 AF2037 hypothetical prote
362 30 53.6 219 1 TRPGAZ azurocidin - pig
363 30 53.6 227 2 B89057 protein K09H11.2 [i
364 30 53.6 231 2 G70000 two-component resp
365 30 53.6 234 2 T26560 hypothetical prote
366 30 53.6 235 2 AF2319 hypothetical prote
367 30 53.6 238 2 C86656 glycerol uptake fa
368 30 53.6 238 2 S44969 lmbv protein - Str
369 30 53.6 246 2 AI1717 hypothetical prote
370 30 53.6 249 2 F83477 hypothetical prote
371 30 53.6 255 2 T21686 hypothetical prote
372 30 53.6 257 2 D87516 conserved hypothet
373 30 53.6 262 1 A31372 granzyme A (EC 3.4
374 30 53.6 263 2 D83289 probable enoyl-CoA
375 30 53.6 264 2 I39696 beta-lactamase (EC
376 30 53.6 266 2 C72770 probable DNA-(apur
377 30 53.6 266 2 AI2762 conserved hypothet
378 30 53.6 267 2 E70837 hypothetical prote
379 30 53.6 271 2 H97543 hypothetical prote
380 30 53.6 273 2 S48511 MPT4 protein - yea
381 30 53.6 274 2 D97653 hypothetical prote
382 30 53.6 274 2 AB2877 conserved hypothet
383 30 53.6 275 2 G87465 mazg protein limpo
384 30 53.6 276 2 T45275 oxidoreductase of
385 30 53.6 281 2 D95395 probable Beta lact
386 30 53.6 283 2 E49348 succinoglycan bios
387 30 53.6 289 2 S73935 hypothetical prote
388 30 53.6 290 2 F69897 transcription regu
389 30 53.6 293 2 F64558 conserved hypothet
390 30 53.6 293 2 H71949 hypothetical prote
391 30 53.6 302 2 AH0191 probable polysacch
392 30 53.6 303 2 T13598 trypsin homolog -
393 30 53.6 303 2 B47089 probable AraC-type
394 30 53.6 308 2 E98259 probable transcrip

395 30 53.6 308 2 AD3025 transcription regu
396 30 53.6 312 2 T06473 probable obtusifol
397 30 53.6 315 2 T45928 reductase-like pro
398 30 53.6 316 2 G82557 hypothetical prote
399 30 53.6 319 2 T01822 hypothetical prote
400 30 53.6 323 2 D72661 hypothetical prote
401 30 53.6 328 2 G95392 probable protease
402 30 53.6 334 2 A83662 transcription regu
403 30 53.6 334 2 D86455 T9L6.3 protein - A
404 30 53.6 334 2 AE0154 probable nucleoid-
405 30 53.6 335 2 AC0786 nucleoid-associated
406 30 53.6 335 2 A64988 hypothetical 37.8
407 30 53.6 335 2 H85857 nucleoid-associate
408 30 53.6 335 2 F91013 nucleoid-associate
409 30 53.6 338 2 S67158 26S proteasome reg
410 30 53.6 338 2 H83228 hypothetical prote
411 30 53.6 343 2 F72319 translation initia
412 30 53.6 345 2 AG0366 thiosulfate-bindin
413 30 53.6 347 2 AD2711 aldo/keto reductas
414 30 53.6 355 2 AE0518 hypothetical prote
415 30 53.6 356 2 E84495 En/Spm-like transp
416 30 53.6 360 2 A86175 hypothetical prote
417 30 53.6 362 2 AC3192 hypothetical prote
418 30 53.6 367 2 F86315 hypothetical prote
419 30 53.6 370 2 B97493 hypothetical prote
420 30 53.6 371 2 A86264 protein F3F19.3 [i
421 30 53.6 372 2 T14193 L-ascorbate peroxi
422 30 53.6 372 2 G64064 outer membrane int
423 30 53.6 373 2 T47115 probable 4-carboxy
424 30 53.6 376 2 T34734 hypothetical prote
425 30 53.6 376 2 AD2585 DNA replication an
426 30 53.6 376 2 C97367 recF protein limpo
427 30 53.6 385 2 G87519 hypothetical prote
428 30 53.6 387 2 S32934 aminotransferase p
429 30 53.6 389 2 T47323 hypothetical prote
430 30 53.6 392 1 A24608 gastricsin (EC 3.4
431 30 53.6 396 2 AF0677 probable membrane
432 30 53.6 398 2 H87485 lipid-A-disacchari
433 30 53.6 398 2 B70415 proton/sodium-glut
434 30 53.6 402 2 A10323 transposase, IS285
435 30 53.6 402 2 AB0193 transposase, IS285
436 30 53.6 402 2 T14710 probable transposa
437 30 53.6 402 2 AE0267 transposase, IS285
438 30 53.6 402 2 A10256 transposase, IS285
439 30 53.6 402 2 AH0298 transposase, IS285
440 30 53.6 402 2 AG0264 transposase, IS285
441 30 53.6 402 2 AB0472 transposase, IS285
442 30 53.6 402 2 AH0197 transposase, IS285
443 30 53.6 402 2 AE0190 transposase, IS285
444 30 53.6 402 2 AG0210 transposase, IS285
445 30 53.6 402 2 AG0339 transposase, IS285
446 30 53.6 402 2 AB0358 transposase, IS285
447 30 53.6 402 2 A10181 transposase [impor
448 30 53.6 402 2 AH0193 transposase, IS285
449 30 53.6 402 2 AE0242 transposase, IS285
450 30 53.6 402 2 AG0260 transposase, IS285
451 30 53.6 402 2 AC0341 transposase, IS285
452 30 53.6 402 2 AD0002 transposase, IS285
453 30 53.6 402 2 A10313 transposase, IS285
454 30 53.6 402 2 AF0348 transposase, IS285
455 30 53.6 406 2 T47674 hypothetical prote
456 30 53.6 407 2 G82040 general secretion
457 30 53.6 408 2 B87436 conserved hypothet
458 30 53.6 408 2 C96502 oxalate/formate an
459 30 53.6 410 2 C84176 conserved hypothet
460 30 53.6 412 2 B72391 alanine-glyoxylate
461 30 53.6 414 2 S43253 L-ascorbate peroxi
462 30 53.6 421 2 T10190 L-ascorbate peroxi
463 30 53.6 421 2 S71331 hypothetical prote
464 30 53.6 423 2 T01319 hypothetical prote
465 30 53.6 423 2 S48817 hypothetical prote
466 30 53.6 425 2 F70608 hypothetical prote
467 30 53.6 426 2 C96804 hypothetical prote

468 30 53.6 427 2 H87571
469 30 53.6 428 2 T03251
470 30 53.6 433 2 A84619
471 30 53.6 433 2 B90288
472 30 53.6 433 2 H90359
473 30 53.6 435 1 T07310
474 30 53.6 436 2 B86469
475 30 53.6 437 2 G84505
476 30 53.6 439 2 T28841
477 30 53.6 444 2 D86283
478 30 53.6 445 2 A49189
479 30 53.6 446 2 F84506
480 30 53.6 446 2 C75087
481 30 53.6 453 2 T06475
482 30 53.6 457 2 I51660
483 30 53.6 459 2 G84752
484 30 53.6 466 1 YTB5MA
485 30 53.6 468 2 D84503
486 30 53.6 473 2 T05306
487 30 53.6 476 2 C85041
488 30 53.6 488 2 G81295
489 30 53.6 493 2 A45737
490 30 53.6 502 2 AF3222
491 30 53.6 517 2 E95959
492 30 53.6 520 1 O4BOM
493 30 53.6 520 1 S03188
494 30 53.6 531 2 C95338
495 30 53.6 538 2 A70836
496 30 53.6 538 2 T27433
497 30 53.6 538 2 H83466
498 30 53.6 541 2 B87532
499 30 53.6 544 2 B83347
500 30 53.6 550 2 B70583
501 30 53.6 552 2 E87226
502 30 53.6 570 2 D96766
503 30 53.6 575 1 S52819
504 30 53.6 577 2 E90044
505 30 53.6 578 2 T33916
506 30 53.6 578 2 AG0119
507 30 53.6 578 2 AC0179
508 30 53.6 590 2 T45820
509 30 53.6 591 1 S30145
510 30 53.6 591 2 T45681
511 30 53.6 594 2 I48771
512 30 53.6 595 1 S17180
513 30 53.6 596 2 C44461
514 30 53.6 596 2 G75457
515 30 53.6 597 2 I52859
516 30 53.6 597 2 S18304
517 30 53.6 604 1 HMNZCD
518 30 53.6 607 2 S49017
519 30 53.6 615 1 KFHU12
520 30 53.6 629 2 F86351
521 30 53.6 651 2 T10219
522 30 53.6 653 2 D95848
523 30 53.6 664 2 F83376
524 30 53.6 666 2 T35864
525 30 53.6 671 2 H91045
526 30 53.6 671 2 D85890
527 30 53.6 671 2 A65023
528 30 53.6 674 2 F82844
529 30 53.6 680 2 H84774
530 30 53.6 682 1 S22700
531 30 53.6 683 2 A83674
532 30 53.6 689 2 E69135
533 30 53.6 694 2 T01504
534 30 53.6 711 2 T12525
535 30 53.6 712 2 E83067
536 30 53.6 724 2 C49423
537 30 53.6 773 2 I39597
538 30 53.6 777 2 T44597
539 30 53.6 778 2 F84237
540 30 53.6 778 2 T48897

N-carbamyl-L-amino
calnexin - maize (
probable serine ca
thiamin biosynthes
thiamin biosynthes
protochlorophyllid
protein F12K21.16
probable PttA2-lik
hypothetical prote
T15D22.4 protein -
cholesterol monoox
En/Spm-like transp
hydroxymethylpyrim
probable obtusifol
dopamine D1B recep
En/Spm-like transp
methylenomycin A r
probable serine ca
homeobox protein A
probable DNA-bindin
cytochrome-c oxida
phenylcarbamate hy
conserved hypothet
probable membrane-
cholesterol monoox
cholesterol monoox
hypothetical prote
hypothetical prote
hypothetical prote
probable ATP-bindin
transporter, proba
hypothetical prote
probable transposa
conserved membrane
protein glucosyltr
isocitrate lyase (
hypothetical prote
hypothetical prote
probable OmpA-fami
probable exported
hypothetical prote
ketol-acid reducto
ketol-acid reducto
Slp(w7) - mouse (f
ketol-acid reducto
voltage-dependent
tetraatricopeptide
L-type voltage-gat
calcium channel pr
hemagglutinin - ca
hemagglutinin - ca
coagulation factor
protein T26F17.2 [
protein kinase hom
probable elongatio
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
oligopeptidase A X
probable homeodoma
amphiphysin - chic
hypothetical prote
coenzyme F420-redu
hypothetical prote
hypothetical prote
probable chemotaxi
semaphorin II prec
hypothetical prote
transducer protein
Htr4 transducer [i
transducer protein

541	30	53.6	791	2	AC0017	conserved hypothet	614	30	53.6	1755	2	S69839	TyB protein - yeas
542	30	53.6	822	2	C97211	uncharacterized co	615	30	53.6	1755	2	S69845	TyB protein - yeas
543	30	53.6	861	2	T36381	probable large ATP	616	30	53.6	1755	2	S69980	TyB protein - yeas
544	30	53.6	875	2	T27495	hypothetical prote	617	30	53.6	1756	2	S45867	TyB protein - yeas
545	30	53.6	882	2	AF3036	nitrate reductase	618	30	53.6	1756	2	S69983	TyB protein - yeas
546	30	53.6	892	2	F87325	hypothetical prote	619	30	53.6	1770	2	S69953	TyB protein - yeas
547	30	53.6	899	2	T35299	probable pyruvate	620	30	53.6	1770	2	S58651	TyB protein - yeas
548	30	53.6	903	2	E98249	nitrate reductase	621	30	53.6	1770	2	S70233	TyB protein - yeas
549	30	53.6	906	1	RNBPK1	DNA-directed RNA p	622	30	53.6	1770	2	S69948	TyB protein - yeas
550	30	53.6	911	2	T08105	nitrate reductase	623	30	53.6	1770	2	S69966	TyB protein - yeas
551	30	53.6	911	2	T08108	nitrate reductase	624	30	53.6	1770	2	S45842	TyB protein - yeas
552	30	53.6	917	2	E96807	nitrate reductase	625	30	53.6	1770	2	S70230	TyB protein - yeas
553	30	53.6	917	2	S35228	nitrate reductase	626	30	53.6	1770	2	S69950	TyB protein - yeas
554	30	53.6	919	2	S37786	hypothetical prote	627	30	53.6	1771	2	S53592	TyB protein - yeas
555	30	53.6	976	2	G87389	TonB-dependent rec	628	30	53.6	1793	2	S52601	TyB protein - yeas
556	30	53.6	1026	1	TLBP74	tail fiber protein	629	30	53.6	1810	2	S69973	TyB protein - yeas
557	30	53.6	1054	2	H69377	reverse gyrase (to	630	30	53.6	1968	1	PN0093	genome polyprotein
558	30	53.6	1063	2	T18255	cytoskeleton assem	631	30	53.6	1997	1	S12050	protein-tyrosine-p
559	30	53.6	1064	2	E72729	probable isoleucyl	632	30	53.6	2163	2	T51397	hypothetical prote
560	30	53.6	1078	2	T30879	dynein heavy chain	633	30	53.6	2264	1	GNVVTB	genome polyprotein
561	30	53.6	1078	2	T00494	RNA-directed RNA p	634	30	53.6	2948	2	T22664	hypothetical prote
562	30	53.6	1082	2	T13410	hypothetical prote	635	30	53.6	3450	2	T26963	hypothetical prote
563	30	53.6	1124	2	F71719	hypothetical prote	636	30	53.6	3461	2	T26964	hypothetical prote
564	30	53.6	1134	2	T04587	hypothetical prote	637	30	53.6	4151	2	G70944	probable polyketid
565	30	53.6	1164	2	T06144	disease resistance	638	30	53.6	4464	2	D87755	protein T21E12.4 [
566	30	53.6	1196	2	S40908	TyB protein - yeas	639	29.5	52.7	121	2	T42138	type II secretion
567	30	53.6	1198	2	D96723	hypothetical prote	640	29.5	52.7	141	2	AF3317	hypothetical prote
568	30	53.6	1209	2	E90994	probable regulator	641	29.5	52.7	161	2	T35350	hypothetical prote
569	30	53.6	1209	2	H85839	probable regulator	642	29.5	52.7	251	2	G75440	conserved hypothet
570	30	53.6	1210	2	E64979	hypothetical i38.1	643	29.5	52.7	315	2	S58171	bifunctional cycla
571	30	53.6	1212	2	F83153	probable two-compo	644	29.5	52.7	341	2	E97181	queuine synthetase
572	30	53.6	1243	2	S07278	tail fiber protein	645	29.5	52.7	356	2	AI1954	hypothetical prote
573	30	53.6	1278	2	B22671	TyB protein - yeas	646	29.5	52.7	383	2	T48870	muconate cycloisom
574	30	53.6	1282	2	T30577	DNA topoisomerase	647	29.5	52.7	385	2	T46824	muconate cycloisom
575	30	53.6	1328	2	B22999	TyB protein - yeas	648	29.5	52.7	646	2	A70903	probable excinucle
576	30	53.6	1328	2	S52481	TyB protein - yeas	649	29.5	52.7	647	2	B86979	probable excinucle
577	30	53.6	1328	2	S50948	TyB protein - yeas	650	29.5	52.7	728	2	T36026	ABC excision nucle
578	30	53.6	1328	2	S52894	TyB protein - yeas	651	29.5	52.7	788	2	AE3153	NADP-dependent ald
579	30	53.6	1328	2	B28097	TyB protein - yeas	652	29.5	52.7	798	2	F98134	hypothetical prote
580	30	53.6	1348	2	B23496	TyB protein - yeas	653	29.5	52.7	3433	1	GNWVKV	probable non-ribos
581	30	53.6	1375	2	T13822	frazzled gene prot	654	29.5	52.7	5149	2	F83345	hypothetical prote
582	30	53.6	1376	2	F72609	probable reverse g	655	29	51.8	50	2	H97798	rubredoxin PA5351
583	30	53.6	1526	2	T13823	frazzled gene prot	656	29	51.8	55	2	A82977	hypothetical prote
584	30	53.6	1538	2	B90924	probable ATP-depen	657	29	51.8	67	2	S20172	hypothetical prote
585	30	53.6	1538	2	F85772	ATP-dependent heli	658	29	51.8	73	2	G82382	hypothetical prote
586	30	53.6	1538	2	G64922	probable ATP-depen	659	29	51.8	74	2	B72512	probable ribosomal
587	30	53.6	1583	2	T14176	probable phosphati	660	29	51.8	74	2	C81133	hypothetical prote
588	30	53.6	1598	2	S69967	TyB protein - yeas	661	29	51.8	81	2	S77318	secretory protein
589	30	53.6	1600	2	AB3281	glutamate dehydrog	662	29	51.8	93	2	D82474	hypothetical prote
590	30	53.6	1690	2	T31670	DNA-directed RNA p	663	29	51.8	102	2	C96034	conserved hypothet
591	30	53.6	1735	2	S54784	sex-limited protei	664	29	51.8	110	2	D90328	hypothetical prote
592	30	53.6	1736	2	A29176	sex-limited protei	665	29	51.8	121	2	E90019	30S ribosomal prot
593	30	53.6	1738	1	A24558	complement C4 prec	666	29	51.8	126	2	S53585	hypothetical prote
594	30	53.6	1749	2	S69972	TyB protein - yeas	667	29	51.8	128	2	S72600	hypothetical prote
595	30	53.6	1755	2	S69975	TyB protein - yeas	668	29	51.8	130	2	S40368	Ig kappa chain - h
596	30	53.6	1755	2	S69866	TyB protein - yeas	669	29	51.8	131	2	C72680	hypothetical prote
597	30	53.6	1755	2	S69866	TyB protein - yeas	670	29	51.8	133	2	H75386	hypothetical prote
598	30	53.6	1755	2	S69955	TyB protein - yeas	671	29	51.8	137	2	AB3157	conserved hypothet
599	30	53.6	1755	2	S50663	TyB protein - yeas	672	29	51.8	137	2	G98130	hypothetical prote
600	30	53.6	1755	2	S50641	TyB protein YER138	673	29	51.8	146	2	T36436	probable gntR-fami
601	30	53.6	1755	2	S57047	TyB protein - yeas	674	29	51.8	148	2	F71367	probable PTS syste
602	30	53.6	1755	2	S69951	TyB protein - yeas	675	29	51.8	154	2	AD1037	probable capsid co
603	30	53.6	1755	2	S45736	TyB protein - yeas	676	29	51.8	158	2	AH1902	hypothetical prote
604	30	53.6	1755	2	S69838	TyB protein - yeas	677	29	51.8	164	2	A87556	hypothetical prote
605	30	53.6	1755	2	S61763	TyB protein - yeas	678	29	51.8	165	2	C97979	hypothetical prote
606	30	53.6	1755	2	S57045	TyB protein - yeas	679	29	51.8	178	2	F90168	hypothetical prote
607	30	53.6	1755	2	S70298	TyB protein - yeas	680	29	51.8	179	2	F96835	unknown protein F5
608	30	53.6	1755	2	S69982	TyB protein - yeas	681	29	51.8	179	2	S39078	auxin-induced prot
609	30	53.6	1755	2	S69960	TyB protein - yeas	682	29	51.8	181	1	TISYB	trypsin inhibitor
610	30	53.6	1755	2	S69969	TyB protein - yeas	683	29	51.8	185	2	T10677	pathogenesis-relat
611	30	53.6	1755	2	S69949	TyB protein - yeas	684	29	51.8	185	2	A70682	hypothetical prote
612	30	53.6	1755	2	S69957	TyB protein - yeas	685	29	51.8	188	2	T10859	auxin-induced prot
613	30	53.6	1755	2	S69963	TyB protein - yeas	686	29	51.8	193	2	H95959	hypothetical membr

687 29 51.8 194 2 T10939 auxin-induced prot
688 29 51.8 195 2 E95110 ABC transporter, A
689 29 51.8 199 2 H87416 phosphoglycerate m
690 29 51.8 200 2 T28839 hypothetical prote
691 29 51.8 201 2 T08333 conserved hypotet
692 29 51.8 203 2 JQ1091 trypsin inhibitor
693 29 51.8 204 2 T41513 probable proteasom
694 29 51.8 204 2 JQ1092 trypsin inhibitor
695 29 51.8 205 2 S29251 hypothetical prote
696 29 51.8 206 2 AB2759 conserved hypotet
697 29 51.8 206 2 AB2546 hypothetical prote
698 29 51.8 209 2 S54133 hypothetical prote
699 29 51.8 211 2 T17228 hypothetical prote
700 29 51.8 214 2 B40296 hypothetical prote
701 29 51.8 215 2 B83429 conserved hypotet
702 29 51.8 216 2 AF0160 probable hydrolase
703 29 51.8 217 2 S19190 trypsin inhibitor
704 29 51.8 218 2 T35174 hypothetical prote
705 29 51.8 221 2 D86600 SET domain protein
706 29 51.8 221 2 B81514 conserved hypotet
707 29 51.8 221 2 D72024 set domain protein
708 29 51.8 223 2 S77461 hypothetical prote
709 29 51.8 226 2 D83249 probable hydrolase
710 29 51.8 226 2 H90233 dTDP-4-dehydroram
711 29 51.8 227 2 T45388 hypothetical prote
712 29 51.8 229 2 F84370 hypothetical prote
713 29 51.8 230 2 D75518 probable biopolyme
714 29 51.8 236 2 E84255 hypothetical prote
715 29 51.8 238 2 G87267 ribonuclease PH [i
716 29 51.8 238 2 E70524 hypothetical prote
717 29 51.8 239 2 B83296 probable transcrip
718 29 51.8 242 2 AE2681 conserved hypotet
719 29 51.8 246 2 H97539 hypothetical prote
720 29 51.8 251 2 S23821 hypothetical prote
721 29 51.8 251 2 C45557 regulatory protein
722 29 51.8 251 2 F83344 probable ATP-bind
723 29 51.8 253 2 T41139 proteasome compone
724 29 51.8 257 1 B64042 formate dehydrogen
725 29 51.8 257 2 S56108 26S proteasome reg
726 29 51.8 257 2 B86987 conserved hypotet
727 29 51.8 258 2 B36914 probable nicotinat
728 29 51.8 258 2 D95910 probable membrane-
729 29 51.8 259 2 A87623 PAP2 homolog prote
730 29 51.8 259 2 A96740 F14023.14 importe
731 29 51.8 260 2 T01837 hypothetical prote
732 29 51.8 263 1 ASLJFP vif protein - feli
733 29 51.8 264 2 F96807 unknown protein T3
734 29 51.8 264 2 JU0328 cellulase (EC 3.2.
735 29 51.8 264 2 C70221 conserved hypotet
736 29 51.8 267 2 H96668 protein FIN19.9 [i
737 29 51.8 272 1 C64009 ribonuclease homol
738 29 51.8 274 2 C85084 hypothetical prote
739 29 51.8 275 2 C70924 conserved hypotet
740 29 51.8 276 2 T08178 assemblin, striate
741 29 51.8 280 2 B95931 probable amino aci
742 29 51.8 283 2 T00581 hypothetical prote
743 29 51.8 285 2 A87350 conserved hypotet
744 29 51.8 290 2 I64051 phosphatidylserine
745 29 51.8 291 2 G82136 hypothetical prote
746 29 51.8 293 2 T43586 secretion control
747 29 51.8 293 2 S15320 yopN protein - Yer
748 29 51.8 293 2 T40537 glycine betaine AB
749 29 51.8 294 2 T31946 hypothetical prote
750 29 51.8 297 2 C87535 hypothetical prote
751 29 51.8 298 2 G72531 hypothetical prote
752 29 51.8 302 2 G70778 [acyl-carrier-prot
753 29 51.8 303 2 E95374 probable ABC trans
754 29 51.8 304 2 C97463 hypothetical prote
755 29 51.8 304 2 F84169 hypothetical prote
756 29 51.8 309 2 I53384 4-1BB ligand - mou
757 29 51.8 311 2 T52312 deacetoxycephalosp
758 29 51.8 311 2 A32043 deacetoxycephalosp

760 29 51.8 312 2 S59519
761 29 51.8 319 2 S62196
762 29 51.8 324 2 D95869
763 29 51.8 326 2 I51675
764 29 51.8 327 2 T21873
765 29 51.8 330 2 A86921
766 29 51.8 330 2 D95362
767 29 51.8 331 2 S76480
768 29 51.8 338 1 MMBEK1
769 29 51.8 338 1 MMBEK2
770 29 51.8 338 2 S10326
771 29 51.8 338 2 H70887
772 29 51.8 339 2 S76619
773 29 51.8 341 2 F86483
774 29 51.8 341 2 F87620
775 29 51.8 346 2 D64484
776 29 51.8 347 2 H82391
777 29 51.8 349 2 H72715
778 29 51.8 350 2 T50581
779 29 51.8 351 2 AH2205
780 29 51.8 352 2 D90390
781 29 51.8 362 2 T17477
782 29 51.8 365 2 C86647
783 29 51.8 366 2 A99375
784 29 51.8 366 2 S74005
785 29 51.8 370 2 S75625
786 29 51.8 370 2 C70747
787 29 51.8 371 2 H87374
788 29 51.8 371 2 AI3558
789 29 51.8 372 2 AI2862
790 29 51.8 375 2 S76663
791 29 51.8 376 2 G83290
792 29 51.8 379 2 AB0959
793 29 51.8 381 2 F75569
794 29 51.8 381 2 D70669
795 29 51.8 382 2 C86925
796 29 51.8 383 2 C95965
797 29 51.8 385 2 S44232
798 29 51.8 386 2 H75287
799 29 51.8 386 2 AE1996
800 29 51.8 386 2 T36425
801 29 51.8 388 2 H72223
802 29 51.8 389 2 E81806
803 29 51.8 390 2 H87345
804 29 51.8 390 2 AB3384
805 29 51.8 392 2 G85076
806 29 51.8 393 1 C70929
807 29 51.8 394 2 JC5197
808 29 51.8 394 2 F72517
809 29 51.8 396 1 C64907
810 29 51.8 396 2 F83130
811 29 51.8 396 2 H85721
812 29 51.8 396 2 G90895
813 29 51.8 399 2 B82197
814 29 51.8 399 2 A43864
815 29 51.8 400 2 B70806
816 29 51.8 401 2 B82868
817 29 51.8 403 2 S76603
818 29 51.8 409 2 G86257
819 29 51.8 410 2 H97639
820 29 51.8 411 2 AB2214
821 29 51.8 411 2 E71873
822 29 51.8 412 2 G70735
823 29 51.8 412 2 AC2152
824 29 51.8 413 2 T31663
825 29 51.8 414 2 AG2321
826 29 51.8 414 2 S06729
827 29 51.8 418 2 T15142
828 29 51.8 419 2 S54102
829 29 51.8 420 2 AI0921
830 29 51.8 420 2 S12484
831 29 51.8 428 2 JC4601
832 29 51.8 430 2 B84619

tryptophan synthas
hypothetical prote
probable nutrient
ribonucleoprotein
hypothetical prote
antigen 85A, mycol
probable AraC-fami
hypothetical prote
cell fusion protei
cell fusion protei
alpha-antigen A, e
32K antigen fbpA p
hypothetical prote
unknown protein, 4
conserved hypotet
hypothetical prote
NADH-dependent fla
probable aspartate
binding-protein-de
hypothetical prote
conserved hypotet
probable 2-dehydro
hypothetical prote
transport protein,
conserved hypotet
hypothetical prote
hypothetical prote
perosamine synthet
glycerol trinitrat
oxidoreductase Atu
hypothetical prote
hypothetical prote
2-oxo acid dehydro
coenzyme F420-depe
probable oxidoredu
probable two-compo
strH protein - Str
hypothetical prote
hypothetical prote
probable sugar iso
thiamin biosynthes
O-succinylhomoseri
thiolase family pr
N-acetylmuramoyl-L
probable reverse t
cytochrome P450 Rv
aromatic-amino-aci
hypothetical prote
chloramphenicol re
probable MFS trans
probable resistanc
hypothetical prote
zona occludens tox
zonula occludens t
probable acyl-CoA
plasmid replicatio
hypothetical prote
hypothetical prote
probable D-amino a
glutamate-1-semial
hypothetical prote
hypothetical prote
dolichyl-phosphate
isoleucine-tRNA li
dolichol-phosphate
probable IS117 tra
hypothetical prote
isopenicillin N ep
UDP-ManNAc dehydro
transforming prote
hypothetical 48.2k
probable serine ca

833	29	51.8	430	2	2	B70518	probable cina - My
834	29	51.8	433	2	2	H82075	pho4 family protei
835	29	51.8	434	2	2	G70011	conserved hypothet
836	29	51.8	435	2	2	G75366	glucose-1-phosphat
837	29	51.8	437	2	2	D84619	probable serine ca
838	29	51.8	438	2	2	T35355	hypothetical prote
839	29	51.8	441	2	2	D84511	En/Spm-like transp
840	29	51.8	441	2	2	H82642	hypothetical prote
841	29	51.8	442	2	2	B87523	argininosuccinate
842	29	51.8	442	2	2	T02620	hypothetical prote
843	29	51.8	442	2	2	D95889	conserved hypothet
844	29	51.8	445	2	2	F75389	probable carboxyl-
845	29	51.8	446	2	2	F82609	aminopeptidase P X
846	29	51.8	446	2	2	AG3284	high-affinity bran
847	29	51.8	447	2	2	G96569	hypothetical prote
848	29	51.8	447	2	2	AF3459	hypothetical prote
849	29	51.8	447	2	2	AI3364	hypothetical prote
850	29	51.8	449	2	2	F84241	probable membrane
851	29	51.8	450	2	2	AG0008	hypothetical prote
852	29	51.8	451	2	2	B96495	hypothetical prote
853	29	51.8	453	2	2	E69755	hypothetical prote
854	29	51.8	455	2	2	H84955	UDP-N-acetylmuramo
855	29	51.8	456	2	2	A70920	probable adrenodox
856	29	51.8	463	2	2	G84901	hypothetical prote
857	29	51.8	464	2	2	T17332	hypothetical prote
858	29	51.8	466	2	2	AE2999	argininosuccinate
859	29	51.8	467	2	2	D83055	poly(A) polymerase
860	29	51.8	469	1	1	NMIV27	exo-alpha-sialidas
861	29	51.8	469	1	1	NMIVH3	exo-alpha-sialidas
862	29	51.8	469	1	1	NMIVN2	exo-alpha-sialidas
863	29	51.8	469	1	1	NMIV2	exo-alpha-sialidas
864	29	51.8	469	2	2	JQ1644	exo-alpha-sialidas
865	29	51.8	469	2	2	D82976	hypothetical prote
866	29	51.8	479	2	2	C75513	probable ferredoxi
867	29	51.8	480	2	2	S56639	ribosomal protein
868	29	51.8	481	2	2	B84700	probable flavonol
869	29	51.8	483	2	2	C82090	ribosomal protein
870	29	51.8	486	2	2	D98284	argininosuccinate
871	29	51.8	490	2	2	B71279	probable serine-ty
872	29	51.8	493	2	2	AI2314	hypothetical prote
873	29	51.8	495	1	1	FJEC	transcription term
874	29	51.8	495	2	2	B91135	transcription term
875	29	51.8	495	2	2	E85980	transcription term
876	29	51.8	496	2	2	H83577	probable transcrip
877	29	51.8	497	2	2	T48730	related to PRP9 pr
878	29	51.8	500	1	1	A53377	transcription term
879	29	51.8	503	2	2	T37119	probable membrane-
880	29	51.8	505	2	2	H95390	probable reverse t
881	29	51.8	506	2	2	B84339	hypothetical prote
882	29	51.8	507	2	2	A25228	protein kinase CDC
883	29	51.8	509	2	2	S49224	glycoprotein precu
884	29	51.8	510	2	2	I56242	lymphoid cell acti
885	29	51.8	511	2	2	S50581	hypothetical prote
886	29	51.8	512	2	2	H84310	cobryic acid synth
887	29	51.8	512	2	2	G85916	hypothetical prote
888	29	51.8	512	2	2	D91072	multidrug resistan
889	29	51.8	512	2	2	G65048	multidrug resistan
890	29	51.8	512	2	2	AC3399	probable colanic b
891	29	51.8	514	2	2	S52771	beta-glucosidase (
892	29	51.8	519	2	2	F81811	conserved hypothet
893	29	51.8	524	2	2	F96784	hypothetical prote
894	29	51.8	525	2	2	H70982	probable fadD7 pro
895	29	51.8	525	2	2	E87076	probable secreted
896	29	51.8	528	2	2	C96564	probable beta-gluc
897	29	51.8	528	2	2	D70968	hypothetical prote
898	29	51.8	530	2	2	JN0597	calnexin-like prot
899	29	51.8	532	2	2	T49873	calnexin homolog -
900	29	51.8	538	2	2	H86329	F6F9.25 protein -
901	29	51.8	548	2	2	D87254	N utilization subs
902	29	51.8	556	2	2	T47552	hypothetical prote
903	29	51.8	561	2	2	T46845	K+-transporting Ar
904	29	51.8	564	2	2	C84188	hypothetical prote
905	29	51.8	573	2	2	B70942	hypothetical prote
574	2	51.8	574	2	2	C83725	endo-beta-1,4-gluc
575	2	51.8	575	2	2	A84782	hypothetical prote
584	2	51.8	584	2	2	D84679	probable homeodoma
592	2	51.8	592	2	2	B83231	probable homeodoma
594	1	51.8	594	1	1	WZBEE2	probable short-cha
598	2	51.8	598	2	2	T02795	protein-serine/thr
600	2	51.8	600	2	2	H81331	probable membrane
600	2	51.8	600	2	2	H98149	excinuclease ABC c
600	2	51.8	600	2	2	AD3138	ATP-binding transp
611	2	51.8	611	2	2	S09142	hypothetical prote
627	2	51.8	627	2	2	T02415	ND5 intron 1 prote
629	2	51.8	629	2	2	AE2497	probable homeodoma
638	2	51.8	638	2	2	D85435	hypothetical prote
638	2	51.8	638	2	2	H82690	BEL1-like homeobox
648	2	51.8	648	2	2	B40727	hypothetical prote
650	2	51.8	650	2	2	AG0733	S-M checkpoint con
655	1	51.8	655	1	1	A46688	probable bacteriop
655	2	51.8	655	2	2	E82712	hepatocyte growth
656	2	51.8	656	2	2	AE0315	peptidyl-prolyl ci
660	2	51.8	660	2	2	AH2348	probable carbohydr
664	2	51.8	664	2	2	A69138	hypothetical prote
666	2	51.8	666	2	2	G82111	DNA-dependent RNA
669	2	51.8	669	2	2	I54205	flagellar hook-ass
677	2	51.8	677	2	2	F64139	galactosylceramida
686	2	51.8	686	2	2	F82971	guanosine-3',5'-bi
690	2	51.8	690	2	2	H87691	probable FMN oxido
690	2	51.8	690	2	2	A84639	penicillin-binding
692	2	51.8	692	2	2	T00025	probable glucosylt
692	2	51.8	692	2	2	G95844	PSD-95 binding pro
695	2	51.8	695	2	2	S62400	probable pyruvate
702	1	51.8	702	1	1	SHECGD	amphiphysin (clone
702	2	51.8	702	2	2	F91194	guanosine-3',5'-bi
702	2	51.8	702	2	2	F86041	(p)ppGpp synthetas
702	2	51.8	702	2	2	AG0005	guanosine-3',5'-bi
703	2	51.8	703	2	2	AG0969	guanosine-3',5'-bi
709	2	51.8	709	2	2	T32089	guanosine-3',5'-bi
718	2	51.8	718	2	2	E82697	hypothetical prote
720	2	51.8	720	2	2	T51007	GTP diphosphokinas
731	2	51.8	731	2	2	JC7701	hypothetical prote
769	2	51.8	769	2	2	G83183	hypothetical prote
781	2	51.8	781	2	2	A43866	ARHGAP9 protein -
783	2	51.8	783	2	2	T45899	probable chemotaxi
787	2	51.8	787	2	2	A75347	neuraminidase - Vi
796	2	51.8	796	2	2	T34805	receptor protein k
807	2	51.8	807	2	2	E82158	GTP pyrophosphokin
814	2	51.8	814	2	2	G02390	hypothetical prote
821	2	51.8	821	2	2	S39983	neuraminidase VC17
822	2	51.8	822	2	2	I38728	disintegrin-like m
840	2	51.8	840	2	2	AG0526	ep88 protein - mou
842	2	51.8	842	2	2	C83177	epidermal growth f
844	1	51.8	844	1	1	ZPECPB	penicillin-binding
844	2	51.8	844	2	2	A85499	probable phosphotr
844	2	51.8	844	2	2	A99648	penicillin-binding
846	2	51.8	846	2	2	S57580	peptidoglycan synt
851	2	51.8	851	2	2	T47495	peptidoglycan synt
868	1	51.8	868	1	1	B33926	penicillin-binding
880	2	51.8	880	2	2	G90163	hypothetical prote
902	2	51.8	902	2	2	AD0587	autotransporter pr
908	2	51.8	908	2	2	A83424	DNA-directed RNA p
911	2	51.8	911	2	2	H97053	DNA-directed RNA p
926	1	51.8	926	1	1	OPPGIT	sensor protein Kdp
957	2	51.8	957	2	2	H69141	assimilatory nitra
967	2	51.8	967	2	2	AH2660	penicillin-binding
967	2	51.8	967	2	2	F97442	iodide peroxidase
968	2	51.8	968	2	2	T45746	hypothetical prote
977	2	51.8	977	2	2	T00014	isoleucyl-tRNA syn
994	2	51.8	994	2	2	A32792	hypothetical prote
994	2	51.8	994	2	2	S24359	DAP-1-alpha protei
1025	2	51.8	1025	2	2	AC0930	Ca2+-transporting
1025	2	51.8	1025	2	2	AE0836	Ca2+-transporting
1034	2	51.8	1034	2	2	AB0551	hypothetical prote
							probable bacteriop
							exonuclease SbcC l

probable cina - My
pho4 family protei
conserved hypothet
glucose-1-phosphat
probable serine ca
hypothetical prote
En/Spm-like transp
hypothetical prote
argininosuccinate
hypothetical prote
conserved hypothet
probable carboxyl-
aminopeptidase P X
high-affinity bran
hypothetical prote
hypothetical prote
hypothetical prote
probable membrane
hypothetical prote
hypothetical prote
UDP-N-acetylmuramo
probable adrenodox
hypothetical prote
hypothetical prote
argininosuccinate
poly(A) polymerase
exo-alpha-sialidas
exo-alpha-sialidas
exo-alpha-sialidas
exo-alpha-sialidas
exo-alpha-sialidas
hypothetical prote
probable ferredoxi
ribosomal protein
probable flavonol
ribosomal protein
argininosuccinate
probable serine-ty
hypothetical prote
transcription term
transcription term
transcription term
probable transcrip
related to PRP9 pr
transcription term
probable membrane-
probable reverse t
hypothetical prote
protein kinase CDC
glycoprotein precu
lymphoid cell acti
hypothetical prote
cobryic acid synth
hypothetical prote
multidrug resistan
multidrug resistan
probable colanic b
beta-glucosidase (
conserved hypothet
hypothetical prote
probable fadD7 pro
probable secreted
probable beta-gluc
hypothetical prote
calnexin-like prot
calnexin homolog -
F6F9.25 protein -
N utilization subs
hypothetical prote
K+-transporting Ar
hypothetical prote
hypothetical prote

833	29	51.8	430	2	2	B70518	probable cina - My
834	29	51.8	433	2	2	H82075	pho4 family protei
835	29	51.8	434	2	2	G70011	conserved hypothet
836	29	51.8	435	2	2	G75366	glucose-1-phosphat
837	29	51.8	437	2	2	D84619	probable serine ca
838	29	51.8	438	2	2	T35355	hypothetical prote
839	29	51.8	441	2	2	D84511	En/Spm-like transp
840	29	51.8	441	2	2	H82642	hypothetical prote
841	29	51.8	442	2	2	B87523	argininosuccinate
842	29	51.8	442	2	2	T02620	hypothetical prote
843	29	51.8	442	2	2	D95889	conserved hypothet
844	29	51.8	445	2	2	F75389	probable carboxyl-
845	29	51.8	446	2	2	F82609	aminopeptidase P X
846	29	51.8	446	2	2	AG3284	high-affinity bran
847	29	51.8	447	2	2	G96569	hypothetical prote
848	29	51.8	447	2	2	AF3459	hypothetical prote
849	29	51.8	447	2	2	AI3364	hypothetical prote
850	29	51.8	449	2	2	F84241	probable membrane
851	29	51.8	450	2	2	AG0008	hypothetical prote
852	29	51.8	451	2	2	B96495	hypothetical prote
853	29	51.8	453	2	2	E69755	hypothetical prote
854	29	51.8	455	2	2	H84955	UDP-N-acetylmuramo
855	29	51.8	456	2	2	A70920	probable adrenodox
856	29	51.8	463	2	2	G84901	hypothetical prote
857	29	51.8	464	2	2	T17332	hypothetical prote
858	29	51.8	466	2	2	AE2999	argininosuccinate
859	29	51.8	467	2	2	D83055	poly(A) polymerase
860	29	51.8	469	1	1	NMIV27	exo-alpha-sialidas</

979 RND multidrug effl 29 51.8 1045 2 D83393
980 hypothetical prote 29 51.8 1046 2 S67786
981 apsB protein - Eme 29 51.8 1051 2 T18302
982 cyclic nucleotide 29 51.8 1054 2 T30901
983 ceoB protein [limp 29 51.8 1064 2 H97657
984 AcrB/AcrD/AcrF fam 29 51.8 1064 2 AH2881
985 hypothetical prote 29 51.8 1113 2 T00271
986 p63 related protei 29 51.8 1114 2 T49517
987 hypothetical prote 29 51.8 1128 2 AH1949
988 related to SREBP c 29 51.8 1154 2 T48829
989 hypothetical prote 29 51.8 1194 2 E96624
990 hypothetical expor 29 51.8 1200 2 C96025
991 protoporphyrin IX 29 51.8 1228 2 AE2397
992 hypothetical prote 29 51.8 1230 2 T04181
993 hypothetical prote 29 51.8 1231 2 C84716
994 protein F5A9.24 [i 29 51.8 1254 2 G86379
995 hypothetical prote 29 51.8 1263 2 T19472
996 myosin heavy chain 29 51.8 1271 2 S37958
997 retrotransposon li 29 51.8 1474 2 B85188
998 cell division prot 29 51.8 1485 2 AG0171
999 ABC transport prot 29 51.8 1490 2 T30550
1000 hypoxia-inducible 29 51.8 1505 2 JC4851

ALIGNMENTS

RESULT 1
S76487
probable copper-transporting ATPase (EC 3.6.1.-) - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 09-Jul-2004
C;Accession: S76487
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76487
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-780 <KAN>
A;Cross-references: UNIPROT:P74512; UNIPARC:UPI00000D35A0; EMBL:D90915; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Function:

A;Description: copper transport coupled with ATP hydrolysis
C;Superfamily: Enterococcus copper-transporting ATPase CopA; ATPase nucleotide-binding d
C;Keywords: ATP; copper transport; hydrolase; ion transport; phosphoprotein; transmembra
F;30-59/Domain: heavy-metal-associated homology <HMA>
F;149-505/Domain: ATPase transduction domain homology <ATT>
F;581-725/Domain: ATPase nucleotide-binding domain homology <ATN>
F;35,38/Binding site: copper (Cys) #status predicted
F;302/Active site: Glu #status predicted
F;460/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 73.2%; Score 41; DB 1; Length 780;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|||||
Db 357 RLADQVAGW 365

RESULT 2
E87684
ATP phosphoribosyltransferase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87684
R;Niernman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <STO>
A;Cross-references: UNIPROT:Q9A2P5; UNIPARC:UPI000012C73E; GB:AE005673; NID:gl3425239; P
C;Genetics:
A;Gene: CC3511

Query Match 69.6%; Score 39; DB 2; Length 320;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
|:|||||
Db 13 GRLKDQVEAW 22

RESULT 3
AD1929
hypothetical protein all0983 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD1929
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD1929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <KUR>
A;Cross-references: UNIPROT:Q8VY68; UNIPARC:UPI00000CDF37; GB:BA000019; PIDN:BAB72940.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0983
C;Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 69.6%; Score 39; DB 2; Length 360;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
|:|||||
Db 265 GHVKDQYTGW 274

RESULT 4
T34101
hypothetical protein, neural-specific - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34101; T14159
R;Bossolasco, M.; Croci, L.; Corradi, A.; Consalez, G.G.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z21477
A;Accession: T34101
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1216 <BOS>
A;Cross-references: UNIPROT:Q9QWI6; UNIPARC:UPI0000029B36; EMBL:U59873; NID:g4097422; PI
R;Croci, L.; Bossolasco, M.; Consalez, G.G.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z17896
A;Accession: T14159
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA

A;Residues: 1-1086,'KL',1089-1216 <CRO>
A;Cross-references: UNIPARC:UPI0000027C2C; EMBL:AF040944; NID:g3098417; PID:g3098418; PI
C;Genetics:
A;Map position: 11
A;Note: P140

Query Match 69.6%; Score 39; DB 2; Length 1216;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
||:||||
Db 104 RMREQVGW 112

RESULT 5
B70725
probable linB protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70725
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felcwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70725
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-300 <COL>
A;Cross-references: UNIPROT:Q50642; UNIPARC:UPI000012926D; GB:Z77724; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: linB
C;Superfamily: tropinesterase

Query Match 67.9%; Score 38; DB 2; Length 300;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 252 GRIRDYVRSW 261

RESULT 6
AB2809
conserved hypothetical protein Atul892 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AB2809
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB2809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <KUR>
A;Cross-references: UNIPROT:Q8UE69; UNIPARC:UPI00000D1CF3; GB:AE008688; PIDN:AAL42888.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul892
A;Map position: circular chromosome

Query Match 66.1%; Score 37; DB 2; Length 63;
Best Local Similarity 50.0%; Pred. No. 5;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 22 GRTLDEIGGW 31

RESULT 7
C84233
hypothetical protein tfeA [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84233
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablec
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84233
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-181 <STO>
A;Cross-references: UNIPROT:Q9HRC7; UNIPARC:UPI0000063708; GB:AE004437; NID:g10580335; I
C;Genetics:
A;Gene: tfeA
C;Superfamily: hypothetical protein MJ0777

Query Match 66.1%; Score 37; DB 2; Length 181;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
||||: ||
Db 78 RVRDEDSGW 86

RESULT 8
S02828
albicidin resistance protein - Klebsiella oxytoca
C;Species: Klebsiella oxytoca
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: S02828
R;Walker, M.J.; Birch, R.G.; Pemberton, J.M.
Mol. Microbiol. 2, 443-454, 1988
A;Title: Cloning and characterization of an albicidin resistance gene from Klebsiella o:
A;Reference number: S02828; MUID:89013885; PMID:2845223
A;Accession: S02828
A;Molecule type: DNA
A;Residues: 1-218 <WAL>
A;Cross-references: UNIPROT:P10488; UNIPARC:UPI00001257BF; EMBL:Y00558; NID:g43785; PID:
A;Note: part of this sequence, including the amino end of the mature protein, was confi.
C;Comment: This protein forms an inactive complex with albicidin.
C;Superfamily: Klebsiella oxytoca albicidin resistance protein
C;Keywords: antibiotic resistance
F;1-218/Product: albicidin resistance protein #status experimental <MAT>

Query Match 66.1%; Score 37; DB 1; Length 218;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 48 GRLRWQPAGW 57

RESULT 9
C95880
probable protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasamid pSymB
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95880

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: C95880
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-421 <KUR>
A;Cross-references: UNIPROT:Q92WM5; UNIPARC:UPI00000CB4AA; GB:ALS91985; PIDN:CAC48707.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB20321
A;Genome: plasmid
C;Superfamily: conserved hypothetical protein H11029

Query Match 66.1%; Score 37; DB 2; Length 421;
Best Local Similarity 70.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db 293 GMVPQQVAGW 302

RESULT 10
A;Title: excinuclease ABC chain C uvrC [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: A12146
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: A12146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-627 <KUR>
A;Cross-references: UNIPROT:Q8YTI9; UNIPARC:UPI0000137ED4; GB:BA000019; PIDN:BAB74427.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: uvrC
C;Superfamily: excinuclease ABC chain C

Query Match 66.1%; Score 37; DB 2; Length 627;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RVRDQVAG 9
Db 239 RIRDQIAG 246

RESULT 11
S36741
probable copper-transporting ATPase (EC 3.6.1.-) pacs - Synechococcus sp.
C;Species: Synechococcus sp.
C;Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 16-Jun-2000
C;Accession: S36741
R;Kanamaru, K.; Kashiwagi, S.; Mizuno, T.
FEBS Lett. 330, 99-104, 1993
A;Title: The cyanobacterium, Synechococcus sp. PCC7942, possesses two distinct genes end

A;Reference number: S36741; MUID:93380581; PMID:8370468
A;Accession: S36741
A;Molecule type: DNA
A;Residues: 1-747 <KAN>
A;Cross-references: UNIPARC:UPI0000012620F; GB:D16437; NID:G435124; PIDN:BAA03907.1; PID:
C;Function:
A;Description: copper transport coupled with ATP hydrolysis
C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
C;Keywords: ATP; copper transport; hydrolase; ion transport; phosphoprotein; transmembr
F;9-38/Domain: heavy-metal-associated homology <HMA>
F;136-480/Domain: ATPase transduction domain homology <ATT>
F;557-699/Domain: ATPase nucleotide-binding domain homology <ATN>
F;14,17/Binding site: copper (Cys) #status predicted
F;289/Active site: Glu #status predicted
F;434/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 66.1%; Score 37; DB 1; Length 747;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 344 RLADQVTGW 352

RESULT 12
AH2543
cation-transporting ATPase alr7635 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2543
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2543
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-753 <KUR>
A;Cross-references: UNIPROT:Q8ZS77; UNIPARC:UPI00000CCDC5; GB:AP003602; PIDN:BAB77278.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7635
A;Genome: plasmid
C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d

Query Match 66.1%; Score 37; DB 2; Length 753;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 345 RLADQVTGW 353

RESULT 13
AE2009
cation-transporting ATPase alr1627 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2009
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2009
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-753 <KUR>

A;Cross-references: UNIPROT:Q8YWI6; UNIPARC:UPI00000CEI59; GB:BA000019; PIDN:BAB77993.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1627
C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d

Query Match 66.1%; Score 37; DB 2; Length 753;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|:|:|:|
Db 345 RLADQVTGW 353

RESULT 14

A45407
collagen alpha 3(IV) chain - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C;Accession: A45407; A43903; A23940
R;Exposito, J.Y.; D'Alessio, M.; Di Liberto, M.; Ramirez, F.
J. Biol. Chem. 268, 5249-5254, 1993
A;Title: Complete primary structure of a sea urchin type IV collagen alpha chain and ana
A;Reference number: A45407; MUID:93186842; PMID:8444899

A;Accession: A45407
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1752 <EXP>
A;Cross-references: UNIPROT:Q26312; UNIPARC:UPI000001773DE
A;Note: sequence extracted from NCBI backbone (NCBIP:126841)
R;Wessel, G.M.; Etkin, M.; Benson, S.
Dev. Biol. 148, 261-272, 1991
A;Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ
A;Reference number: A43903; MUID:92038439; PMID:1936564

A;Accession: A43903
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'P', 633-1537, 'G' <WES>
A;Cross-references: UNIPARC:UPI000007C802; GB:S64572; NID:9238616; PIDN:AAB20270.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBIP:64573)
R;Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986
A;Title: Structure and developmentally regulated expression of a Strongylocentrotus purp
A;Reference number: A23940; MUID:86205894; PMID:3458186

A;Accession: A23940
A;Molecule type: DNA
A;Residues: 742-812 <VEN>
A;Cross-references: UNIPARC:UPI00001773DF; EMBL:M13206
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;29-161/Domain: amino-terminal nonhelical, 7S <7SD>
F;162-1523/Region: interrupted helical
F;1524-1752/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;129/Modified site: allylsine (Lys) #status predicted

Query Match 66.1%; Score 37; DB 2; Length 1752;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
|:|:|:|:|
Db 1165 GIIRDLLAGW 1174

RESULT 15

G82724
hypothetical protein XF1092 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82724
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <SIM>
A;Cross-references: UNIPROT:Q9PED6; UNIPARC:UPI00000C2605; GB:AE003945; GB:AE003849; NII
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; R;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, i
as-Neto, E.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshahako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1092

Query Match 64.3%; Score 36; DB 2; Length 67;
Best Local Similarity 55.6%; Pred. No. 8.3;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|:|:|:|
Db 20 RIRDETASW 28

RESULT 16

G82860
replication protein xFb0001 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82860
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82860
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <SIM>
A;Cross-references: UNIPROT:Q9PHK6; UNIPARC:UPI00000C223E; GB:AE003850; NID:g9112233; P
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvei
M.; Tshahako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XFB0001

A;Genome: plasmid

A;Note: plasmid pXF1.3

C;Superfamily: class II filamentous phage gene 278 protein

Query Match		64.3%;	Score 36;	DB 2;	Length 283;		
Best Local Similarity		60.0%;	Pred. No. 37;				
Matches	6;	Conservative	1;	Mismatches	3; Indels 0; Gaps 0;		
Qy	1	GRVRDQVAGW	10				
		:					
Db	143	GRIIDGVVGW	152				
RESULT 17							
C75353							
ribosomal large subunit pseudouridine synthase D - Deinococcus radiodurans (strain R1)							
C/Species: Deinococcus radiodurans							
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004							
C/Accession: C75353							
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;							
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma							
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.							
Science 286, 1571-1577, 1999							
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.							
A/Reference number: A75250; MUID:20036896; PMID:10567266							
A/Accession: C75353							
A/Status: preliminary							
A/Molecule type: DNA							
A/Residues: 1-338 <WHI>							
A/Cross-references: UNIPROT:Q9RTH5; UNIPARC:UPI00000D3ED4; GB:AE000513; NID							
A/Experimental source: strain R1							
C/Genetics:							
A/Gene: DR1789							
A/Map position: 1							
C/Superfamily: conserved hypothetical protein HI0176							
Query Match		64.3%;	Score 36;	DB 2;	Length 338;		
Best Local Similarity		70.0%;	Pred. No. 44;				
Matches	7;	Conservative	0;	Mismatches	3; Indels 0; Gaps 0;		
Qy	1	GRVRDQVAGW	10				
Db	42	GVSRSQVAGW	51				
RESULT 18							
E83416							
hypothetical protein PA1829 [imported] - Pseudomonas aeruginosa (strain PA01)							
C/Species: Pseudomonas aeruginosa							
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004							
C/Accession: E83416							
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B							
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,							
.; Lory, S.; Olson, M.V.							
Nature 406, 959-964, 2000							
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho							
A/Reference number: A82950; MUID:20437337; PMID:10984043							
A/Accession: E83416							
A/Status: preliminary							
A/Molecule type: DNA							
A/Residues: 1-356 <STO>							
A/Cross-references: UNIPROT:Q9I2R6; UNIPARC:UPI00000C54B9; GB:AE004609; GB:AE004091; NID							
A/Experimental source: strain PA01							
C/Genetics:							
A/Gene: PA1829							
Query Match		64.3%;	Score 36;	DB 2;	Length 356;		
Best Local Similarity		60.0%;	Pred. No. 47;				
Matches	6;	Conservative	2;	Mismatches	2; Indels 0; Gaps 0;		
Qy	1	GRVRDQVAGW	10				
Db	168	GYVQRQIAGW	177				
RESULT 19							
F70686							

hypothetical protein Rv1723 - Mycobacterium tuberculosis (strain H37RV)	
C/Species: Mycobacterium tuberculosis	
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004	
C/Accession: F70686	
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S	
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.	
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.	
Nature 393, 537-544, 1998	
A/Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.	
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome	
A/Reference number: A70500; MUID:98295987; PMID:9634230	
A/Accession: F70686	
A/Status: preliminary; nucleic acid sequence not shown; translation not shown	
A/Molecule type: DNA	
A/Residues: 1-415 <COL>	
A/Cross-references: UNIPROT:P71981; UNIPARC:UPI00000D5F1E; GB:Z81360; GB:AL123456; NID:9	
A/Experimental source: strain H37RV	
C/Genetics:	
A/Gene: Rv1723	
Query Match	
Best Local Similarity	
Matches	
Qy	
Db	
RESULT 20	
G86207	
hypothetical protein [imported] - Arabidopsis thaliana	
C/Species: Arabidopsis thaliana (mouse-ear cress)	
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
C/Accession: G86207	
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,	
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;	
ansen, N.F.; Hughes, B.; Huizar, L.	
Nature 408, 816-820, 2000	
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.	
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,	
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,	
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
A/Reference number: A86141; MUID:21016719; PMID:11130712	
A/Accession: G86207	
A/Status: preliminary	
A/Molecule type: DNA	
A/Residues: 1-479 <STO>	
A/Cross-references: UNIPROT:Q9LML6; UNIPARC:UPI000000BD55; GB:AE005172; NID:g10567858; P	
C/Genetics:	
A/Map position: 1	
C/Superfamily: flavonol O3-glucosyltransferase	
Query Match	
Best Local Similarity	
Matches	
Qy	
Db	
RESULT 21	
I39066	
N-methyl-D-aspartate receptor chain NMDAR2B - human (fragments)	
N/Alternate names: glutamate receptor, ionotropic, N-methyl D-aspartate 2B; NMDA recepto	
C/Species: Homo sapiens (man)	
C/Date: 01-Mar-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004	
C/Accession: I39066; I39067; I39064	
R/Mandich, P.; Schito, A.M.; Bellone, E.; Antonacaci, R.; Finelli, P.; Rocchi, M.; Ajmar,	
Genomics 22, 216-218, 1994	

A;Gene: VC1068
A;Map position: 1

Query Match 62.5%; Score 35; DB 2; Length 113;
Best Local Similarity 55.6%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:|
Db 69 RISDQLPGW 77

RESULT 26
B69047
arsenate reductase - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: B69047
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: B69047
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-130 <MTH>
A;Cross-references: UNIPROT:O27408; UNIPARC:UPI000006656D; GB:AE000898; GB:AE000666; NID
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1355
C;Superfamily: protein-tyrosine-phosphatase, low molecular weight

Query Match 62.5%; Score 35; DB 2; Length 130;
Best Local Similarity 44.4%; Pred. No. 25;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:|
Db 112 RIRDEISDW 120

RESULT 27
A41840
hypothetical protein (hadA 5' region) - Pseudomonas putida (fragment)
C;Species: Pseudomonas putida
C;Date: 04-Mar-1993 #sequence_revision 26-Apr-1996 #text_change 21-Aug-1998
C;Accession: A41840
R;Barth, P.T.; Bolton, L.; Thomson, J.C.
J. Bacteriol. 174, 2612-2619, 1992
A;Title: Cloning and partial sequencing of an operon encoding two Pseudomonas putida hal
A;Reference number: A41840; MUID:92210506; PMID:1556080
A;Contents: AJ1
A;Accession: A41840
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-150 <BAR>
A;Cross-references: UNIPARC:UPI0000178057
A;Note: sequence extracted from NCBI backbone (NCBIN:94091, NCBIP:94092)
C;Superfamily: proline carrier protein

Query Match 62.5%; Score 35; DB 2; Length 150;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:|
Db 118 RVRGHLAGW 126

RESULT 28
AD2083

phosphonate ABC transporter, ATP-binding component phnK [imported] - Nostoc sp. (strain
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004
C;Accession: AD2083
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2083
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-257 <KUR>
A;Cross-references: UNIPROT:Q8YUW3; UNIPARC:UPI00000CE374; GB:BA000019; PIDN:BAB73917.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: phnK

Query Match 62.5%; Score 35; DB 2; Length 257;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|:|:|
Db 124 GNIrDEAAHW 133

RESULT 29
S76855
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76855
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpō, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-300 <KAN>
A;Cross-references: UNIPROT:P74651; UNIPARC:UPI00000C09EB; EMBL:D90917; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 62.5%; Score 35; DB 2; Length 300;
Best Local Similarity 55.6%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:|
Db 206 RYKDEVVGVW 214

RESULT 30
B87128
conserved hypothetical protein ML1752 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87128
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: B87128
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-302 <STO>	
A;Cross-references: UNIPROT:Q9CBP4; UNIPARC:UPI000000C6DE1; GB:AL450380; NID:gl3093491; H	
C;Genetics:	
A;Gene: ML1752	
Query Match 62.5%; Score 35; DB 2; Length 302;	
Best Local Similarity 50.0%; Pred. No. 61;	
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;	
QY	1 GRVRDQVAGW 10
Db	27 GEVKDRSSGW 36
RESULT 31	
G83219	
probable dihydrolipoamide acetyltransferase PA3415 [imported] - Pseudomonas aeruginosa	
C;Species: Pseudomonas aeruginosa	
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004	
C;Accession: G83219	
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu	
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,	
.; Lory, S.; Olson, M.V.	
Nature 406, 959-964, 2000	
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho	
A;Reference number: A82950; MUID:20437337; PMID:10984043	
A;Accession: G83219	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-370 <STO>	
A;Cross-references: UNIPROT:Q9HYJ0; UNIPARC:UPI000000C59FD; GB:AE004762; GB:AE004091; NID	
A;Experimental source: strain PA01	
C;Genetics:	
A;Gene: PA3415	
Query Match 62.5%; Score 35; DB 2; Length 370;	
Best Local Similarity 50.0%; Pred. No. 75;	
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY	1 GRVRDQVAGW 10
Db	319 GGIRDEVVAW 328
RESULT 32	
S38461	
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Chlorobium limicola	
C;Species: Chlorobium limicola	
C;Date: 22-Jan-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004	
C;Accession: S38461	
R;Schuetz, M.; Zirngibl, S.; Le Coutre, J.; Buettner, M.; Xie, D.L.; Nelson, N.; Deutzma	
submitted to the EMBL Data Library, June 1993	
A;Description: A transcription unit for the Rieske FeS-protein and cytochrome b in Chlo	
A;Reference number: S38460	
A;Accession: S38461	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-428 <SCH>	
A;Cross-references: UNIPROT:Q59297; UNIPARC:UPI000016EB31; EMBL:X73628; NID:g410433; PID	
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol	
C;Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase	
F;89-423/Domain: cytochrome b6 homology <CBH>	
F;89-289/Domain: cytochrome b6 homology <CB6>	
F;304-423/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>	
F;161,261/Binding site: heme iron (His) (axial ligands) (low potential) #status predicte	
F;175,276/Binding site: heme iron (His) (axial ligands) (high potential) #status predict	
Query Match 62.5%; Score 35; DB 2; Length 428;	
Best Local Similarity 85.7%; Pred. No. 87;	
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	4 RDQVAGW 10

Db	74 RDAVAGW 80
RESULT 33	
G84823	
probable anthocyanin 5-aromatic acyltransferase [imported] - Arabidopsis thaliana	
C;Species: Arabidopsis thaliana (mouse-ear cress)	
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004	
C;Accession: G84823	
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;	
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L	
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J	
Nature 402, 761-768, 1999	
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.	
A;Reference number: A84420; MUID:20083487; PMID:10617197	
A;Accession: G84823	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-482 <STO>	
A;Cross-references: UNIPROT:O04201; UNIPARC:UPI000000396F7; GB:AE002093; NID:g2088651; P	
C;Genetics:	
A;Gene: At2g39980	
A;Map position: 2	
C;Superfamily: N-hydroxycinnamoyl/benzoyl transferase	
Query Match 62.5%; Score 35; DB 2; Length 482;	
Best Local Similarity 60.0%; Pred. No. 99;	
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 GRVRDQVAGW 10
Db	382 GRIKSVVADW 391
RESULT 34	
E95335	
probable cation transport P-type ATPase (EC 3.6.3.-) [imported] - Sinorhizobium melilot;	
C;Species: Sinorhizobium meliloti	
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004	
C;Accession: E95335	
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow	
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C	
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001	
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo-	
A;Reference number: A95262; MUID:21396509; PMID:11481432	
A;Accession: E95335	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-733 <KUR>	
A;Cross-references: UNIPROT:Q92ZA2; UNIPARC:UPI000000CB13D; GB:AE006469; PIDN:AAK55247.1	
A;Experimental source: strain 1021, megaplasmid pSyma	
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler	
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.	
L.; Hyman, R.W.; Jones, T.	
Science 293, 668-672, 2001	
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure	
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K	
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.	
A;Reference number: A96039; MUID:21368234; PMID:11474104	
A;Contents: annotation	
C;Genetics:	
A;Gene: SMa1087	
A;Genome: plasmid	
C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding	
C;Keywords: hydrolase	
Query Match 62.5%; Score 35; DB 2; Length 733;	
Best Local Similarity 55.6%; Pred. No. 1.5e+02;	
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	2 RVRDQVAGW 10
Db	330 RLAEVSGW 338

RESULT 35
H75507
hypothetical protein ~ Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75507
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, B.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75507
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-891 <WHI>
A;Cross-references: UNIPROT:Q9RWY5; UNIPARC:UPI000000C1789; GB:AE001911; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0530
A;Map position: 1

Query Match 62.5%; Score 35; DB 2; Length 891;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
|||::||
Db 566 VRDELTGW 573

RESULT 36
C72409
reverse gyrase - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: C72409
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: C72409
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1104 <ARN>
A;Cross-references: UNIPROT:O51934; UNIPARC:UPI000000D382F; GB:AE001702; GB:AE000512; NID
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0173

Query Match 62.5%; Score 35; DB 2; Length 1104;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||::||
Db 752 GRVQSTVLGW 761

RESULT 37
E72611
probable ATP-dependent DNA helicase APE1353 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72611
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72611
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1273 <KAW>
A;Cross-references: UNIPROT:Q9YCAL; UNIPARC:UPI000005DEF8; DDBJ:AP0000061; NID:G5104821;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1353

Query Match 62.5%; Score 35; DB 2; Length 1273;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||::||
Db 341 GRLRIRVRGW 350

RESULT 38
AC3434
calf thymus ribonuclease H (EC 3.1.1.26.4) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3434
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3434
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <KUR>
A;Cross-references: UNIPROT:Q8YFR3; UNIPARC:UPI0000058097; GB:AE008917; PIDN:AAL52638.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BME11457
A;Map position: I
C;Superfamily: ribonuclease H

Query Match 60.7%; Score 34; DB 2; Length 154;
Best Local Similarity 62.5%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
|||::||
Db 73 VRDGISGW 80

RESULT 39
C72782
hypothetical protein APE0244 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72782
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72782
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-186 <KAW>
A;Cross-references: UNIPROT:Q9YFK4; UNIPARC:UPI000005DA75; DDBJ:AP0000058; NID:G5103388;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0244
C;Superfamily: conserved hypothetical protein HI1648

Query Match 60.7%; Score 34; DB 2; Length 186;
Best Local Similarity 66.7%; Pred. No. 57;

Matches	6;	Conservative	2;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	GRVRDQVAG	9						
Db	75	GKVRDKVVG	83						
RESULT 40									
D83620									
hypothetical protein PA0201 [imported] - Pseudomonas aeruginosa (strain PA01)									
C;Species: Pseudomonas aeruginosa									
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004									
C;Accession: D83620									
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz									
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,									
.; Lory, S.; Olson, M.V.									
Nature 406, 959-964, 2000									
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho									
A;Reference number: A82950; MUID:20437337; PMID:10984043									
A;Accession: D83620									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-193 <STO>									
A;Cross-references: UNIPROT:Q916T7; UNIPARC:UPI00000C4F91; GB:AE004458; GB:AE004091; NID									
A;Experimental source: strain PA01									
C;Genetics:									
A;Gene: PA0201									
Query Match 60.7%; Score 34; DB 2; Length 193;									
Best Local Similarity 87.5%; Pred. No. 59;									
Matches	7;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	2	RVRDQVAG	9						
Db	86	RVRDQVRG	93						
RESULT 41									
G87457									
guanylate kinase [imported] - Caulobacter crescentus									
C;Species: Caulobacter crescentus									
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004									
C;Accession: G87457									
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.									
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo									
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.									
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001									
A;Title: Complete Genome Sequence of Caulobacter crescentus.									
A;Reference number: A87249; MUID:21173698; PMID:11259647									
A;Accession: G87457									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-213 <STO>									
A;Cross-references: UNIPROT:Q9A7N9; UNIPARC:UPI000012DDD8; GB:AE005673; NID:gl3423091; F									
C;Genetics:									
C;Superfamily: guanylate kinase; guanylate kinase homology									
A;Gene: CC1681									
Query Match 60.7%; Score 34; DB 2; Length 213;									
Best Local Similarity 55.6%; Pred. No. 65;									
Matches	5;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0;
QY	2	RVRDQVAGW	10						
Db	157	RAKDEVAAW	165						
RESULT 42									
C84229									
hypothetical protein Vng0718c [imported] - Halobacterium sp. NRC-1									
C;Species: Halobacterium sp. NRC-1									
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004									
C;Accession: C84229									

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.									
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000									
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li									
A;Title: Genome sequence of Halobacterium species NRC-1.									
A;Reference number: A84160; MUID:20504483; PMID:11016950									
A;Accession: C84229									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-225 <STO>									
A;Cross-references: UNIPROT:Q9HRF9; UNIPARC:UPI00000636ED; GB:AE004437; NID:gl0580297; E									
C;Genetics:									
A;Gene: VNG0718C									
Query Match 60.7%; Score 34; DB 2; Length 225;									
Best Local Similarity 60.0%; Pred. No. 69;									
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;									
QY	1	GRVRDQVAGW	10						
			:						
Db	25	GRVLDVLRGW	34						
RESULT 43									
F87259									
hypothetical protein CC0087 [imported] - Caulobacter crescentus									
C;Species: Caulobacter crescentus									
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004									
C;Accession: F87259									
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J									
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo									
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M									
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001									
A;Title: Complete Genome Sequence of Caulobacter crescentus.									
A;Reference number: A87249; MUID:21173698; PMID:11259647									
A;Accession: F87259									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-252 <STO>									
A;Cross-references: UNIPROT:Q9ABY3; UNIPARC:UPI00000C6F16; GB:AE005673; NID:gl3421188; ,									
C;Genetics:									
A;Gene: CC0087									
Query Match 60.7%; Score 34; DB 2; Length 252;									
Best Local Similarity 71.4%; Pred. No. 78;									
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
QY	4	RDQVAGW	10						
			:						
Db	185	KDQVVGW	191						
RESULT 44									
AI3628									
peptidylprolyl isomerase (EC 5.2.1.8) [imported] - Brucella melitensis (strain 16M)									
C;Species: Brucella melitensis									
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004									
C;Accession: AI3628									
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,									
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes									
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002									
A;Title: The genome sequence of the facultative intracellular pathogen Brucella meliten									
A;Reference number: AD3252; PMID:11756688									
A;Accession: AI3628									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-271 <KUR>									
A;Cross-references: UNIPROT:Q8YBE6; UNIPARC:UPI00000586A7; GB:AE008918; PIDN:AAL54196.1									
A;Experimental source: strain 16M									
C;Genetics:									
A;Gene: BMEI10954									
A;Map position: 11									

C;Keywords: cis-trans-isomerase

Query Match 60.7%; Score 34; DB 2; Length 271;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
| | : | | |
Db 228 VADKIAGW 235

RESULT 45
Tl6323
hypothetical protein F41C6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: Tl6323
R;Geisel, C.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid F41C6.
A;Reference number: Zl8495
A;Accession: Tl6323
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-337 <GEI>
A;Cross-references: UNIPROT:Q20269; UNIPARC:UPI000007AD01; EMBL:U39745; NID:g1049470; PI
C;Genetics:
A;Gene: CESP:F41C6.4
A;Introns: 27/2; 105/2; 219/3

Query Match 60.7%; Score 34; DB 2; Length 337;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
: | : | | |
Db 173 KLRENIAGW 181

RESULT 46
G71499
probable D-Ala-D-Ala carboxypeptidase - Chlamydia trachomatis (serotype D, strain UW3/Cx
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: G71499
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: G71499
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <ARN>
A;Cross-references: UNIPROT:O84555; UNIPARC:UPI00000D335A; GB:AE001326; GB:AE001273; NID
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: dacC

Query Match 60.7%; Score 34; DB 2; Length 343;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
: | : | | |
Db 120 KVKEVSGW 128

RESULT 47
G95297
probable hydrolasee [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymA
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 31-Dec-2004
C;Accession: G95297

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95297
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <KUR>
A;Cross-references: UNIPROT:Q930B1; UNIPARC:UPI00000CB048; GB:AE006469; PIDN:AAK64945.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0551
A;Genome: plasmid
C;Superfamily: metal-dependent hydrolase (amidohydrolase)

Query Match 60.7%; Score 34; DB 2; Length 434;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVRDQVAG 9
| : | | | : | |
Db 354 GQVRDPIAG 362

RESULT 48
C65203
argininosuccinate lyase (EC 4.3.2.1) - Escherichia coli (strain K-12)
N;Alternate names: argininosuccinase
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C65203; JT0333; I41139
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C65203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-457 <BLAT>
A;Cross-references: UNIPROT:P11447; UNIPARC:UPI0000125EF5; GB:AE000470; GB:U00096; NID:g
A;Experimental source: strain K-12, substrain MG1655
R;Parsot, C.; Boyen, A.; Cohen, G.N.; Glansdorff, N.
Gene 68, 275-283, 1988
A;Title: Nucleotide sequence of Escherichia coli argB and argC genes: comparison of N-ac
es.
A;Reference number: JT0331; MUID:89121510; PMID:2851495
A;Accession: JT0333
A;Molecule type: DNA
A;Residues: 1-51 <PAR>
A;Cross-references: UNIPARC:UPI000016EDF0; GB:M21446; NID:g145332; PIDN:AAA23479.1; PID:
R;Charlier, D.; Piette, J.; Glansdorff, N.
Nucleic Acids Res. 10, 5935-5948, 1982
A;Title: i83 can function as a mobile promoter in e.coli.
A;Reference number: I41137; MUID:83064529; PMID:6292860
A;Accession: I41139
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-27,'G',29-39,'R',41-44 <RES>
A;Cross-references: UNIPARC:UPI00000B08F9; GB:J01590; NID:g145350; PIDN:AAB59147.1; PID:
C;Comment: In arginine biosynthesis glutamate is first converted to N-acetylglutamate wh
C;Comment: The latter two reactions are catalyzed by N-acetylglutamatekinase and N-acety
C;Genetics:

Search completed: May 12, 2006, 10:51:04
Job time : 43.2368 secs

A;Gene: argH			
C;Superfamily: argininosuccinate lyase			
C;Keywords: amidine-lyase; carbon-nitrogen lyase			
Query Match 60.7%; Score 34; DB 2; Length 457;			
Best Local Similarity 71.4%; Pred. No. 1.4e+02;			
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
QY	4 RDQVAGW 10		
	.: :		
Db	209 REQLAGW 215		
RESULT 49			
F86087			
argininosuccinate lyase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)			
C;Species: Escherichia coli			
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004			
C;Accession: F86087			
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew			
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,			
Nature 409, 529-533, 2001			
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.			
A;Reference number: A85480; MUID:21074935; PMID:11206551			
A;Accession: F86087			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-457 <STO>			
A;Cross-references: UNIPROT:Q8X730; UNIPARC:UPI0000125EF3; GB:AE005174; NID:g12518882; F			
A;Experimental source: strain O157:H7, substrain EDL933			
C;Genetics:			
A;Gene: argH			
C;Superfamily: argininosuccinate lyase			
Query Match 60.7%; Score 34; DB 2; Length 457;			
Best Local Similarity 71.4%; Pred. No. 1.4e+02;			
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
QY	4 RDQVAGW 10		
	.: :		
Db	209 REQLAGW 215		
RESULT 50			
A98240			
argininosuccinate lyase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05			
C;Species: Escherichia coli			
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004			
C;Accession: A98240			
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.			
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.			
DNA Res. 8, 11-22, 2001			
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno			
A;Reference number: A99629; MUID:21156231; PMID:11258796			
A;Accession: A98240			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-457 <HAY>			
A;Cross-references: UNIPROT:Q8X730; UNIPARC:UPI0000125EF3; GB:BA000007; PIDN:BAB38312.1;			
A;Experimental source: strain O157:H7, substrain RIMD 0509952			
C;Genetics:			
A;Gene: ECs4889			
C;Superfamily: argininosuccinate lyase			
Query Match 60.7%; Score 34; DB 2; Length 457;			
Best Local Similarity 71.4%; Pred. No. 1.4e+02;			
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
QY	4 RDQVAGW 10		
	.: :		
Db	209 REQLAGW 215		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-35
Perfect score: 56
Sequence: 1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB ID	Description
		Match	Length		
1	43	76.8	242	Q4JY45_CORJK	Q4jy45 corynebacte
2	41	73.2	467	Q8RSP4_RHILV	Q8rsp4 rhizobium l
3	41	73.2	776	Q4TM05_9SPHN	Q4tm05 erythrobact
4	41	73.2	780	P74512_SYNY3	P74512 synechocyst
5	41	73.2	823	Q89WI7_BRAJA	Q89wi7 bradyrhizob
6	41	73.2	832	Q9KZY7_STRCO	Q9kzy7 streptomyce
7	40	71.4	130	Q7VLA1_HAEDU	Q7vla1 haemophilus
8	40	71.4	294	Q9RL18_STRCO	Q9rl18 streptomyce
9	40	71.4	302	Q7RL87_PLAYO	Q7rl87 plasmodium
10	40	71.4	313	Q4Z1J9_PLABE	Q4z1j9 plasmodium
11	40	71.4	315	Q4XWR8_PLACH	Q4xwr8 plasmodium
12	40	71.4	320	Q8I2X0_PLAF7	Q8i2x0 plasmodium
13	40	71.4	369	Q7RXX6_NEUCR	Q7rxx6 neurospora
14	39	69.6	98	Q7TLZ4_9ZZZZ	Q7tlz4 uncultured
15	39	69.6	124	Q9FRG1_ORYSA	Q9frg1 oryza sativ
16	39	69.6	320	HIS1 CAUCR	Q9a2p5 caulobacter
17	39	69.6	334	Q4LRX7_9BURK	Q4lrx7 burkholderi
18	39	69.6	360	Q8YY68_ANASP	Q8yy68 anabaena sp
19	39	69.6	1250	SNIP_MOUSE	Q9qwi6 mus musculu
20	39	69.6	1370	Q4QDJ2_LEIMA	Q4qdj2 leishmania
21	39	69.6	1808	Q5LWT4_SILPO	Q5lwt4 silicibacte
22	38	67.9	167	Q8H1Z7_9ROSI	Q8hlz7 citrus gran
23	38	67.9	233	Q62FH7_BURMA	Q62fh7 burkholderi
24	38	67.9	234	Q63VC8_BURPS	Q63vc8 burkholderi
25	38	67.9	243	Q4WRF3_ASPFU	Q4wrf3 aspergillus
26	38	67.9	300	DHAA MYCBO	Q9xb14 mycobacteri
27	38	67.9	300	DHAA MYCTU	Q50642 mycobacteri
28	38	67.9	349	Q7UV53_RHOBA	Q7uv53 rhodopirell
29	38	67.9	367	Q4H705_9DEIO	Q4h705 deinococcus
30	38	67.9	367	Q89SF5_BRAJA	Q89sf5 bradyrhizob
31	38	67.9	369	Q747F0_GEOSL	Q747f0 geobacter s

32	38	67.9	385	2	Q9L144_STRCO	Q9l144 streptomyce
33	38	67.9	443	2	Q74SA4_YERPE	Q74sa4 yersinia pe
34	38	67.9	448	2	Q66HY4_BRARE	Q66hy4 brachydanio
35	38	67.9	461	2	Q668C3_YERPS	Q668c3 yersinia ps
36	38	67.9	588	2	Q7X541_9ACTO	Q7x541 actinoplane
37	38	67.9	941	2	Q869Y0_DICDI	Q869y0 dictyosteli
38	38	67.9	941	2	Q553Y8_DICDI	Q553y8 dictyosteli
39	37	66.1	63	2	Q8UE69_AGRTS	Q8ue69 agrobacteri
40	37	66.1	121	2	Q4LUS0_9BURK	Q4lus0 burkholderi
41	37	66.1	178	2	Q606W1_METCA	Q606w1 methylococc
42	37	66.1	181	2	Q9HRC7_HALSA	Q9hrc7 halobacteri
43	37	66.1	218	1	ALBR_KLEOX	P10488 klebsiella
44	37	66.1	244	2	Q67JS4_SYMTH	Q67js4 symbiobacte
45	37	66.1	276	2	Q5UQ11_MIMIV	Q5uq11 mimivirus.
46	37	66.1	307	2	Q4UPY2_XANCP	Q4upy2 xanthomonas
47	37	66.1	307	2	Q4NPE5_9DELT	Q4npe5 anaeromyxob
48	37	66.1	307	2	Q8P4D1_XANCP	Q8p4d1 xanthomonas
49	37	66.1	335	2	Q4H433_SOLTU	Q4h433 solanum tub
50	37	66.1	372	2	Q50N02_ENTHI	Q50n02 entamoeba h
51	37	66.1	395	2	Q5QUK4_IDILO	Q5quk4 idiomarina
52	37	66.1	420	2	Q8CEB1_MOUSE	Q8ceb1 mus musculu
53	37	66.1	421	2	Q92WM5_RHIME	Q92wm5 rhizobium m
54	37	66.1	431	2	Q7WLQ2_BORBR	Q7wlq2 bordetella
55	37	66.1	461	2	Q6ZE25_ORYSA	Q6ze25 oryza sativ
56	37	66.1	477	2	Q8RTW0_9PROT	Q8rtw0 uncultured
57	37	66.1	477	2	Q6SFG3_9BACT	Q6sfg3 uncultured
58	37	66.1	508	2	Q6DFS1_XENTR	Q6dfs1 xenopus tro
59	37	66.1	508	2	Q6GP74_XENLA	Q6gp74 xenopus lae
60	37	66.1	510	1	ENP1_MOUSE	P55772 mus musculu
61	37	66.1	510	2	Q921Q6_MOUSE	Q921q6 mus musculu
62	37	66.1	510	2	Q544U5_MOUSE	Q544u5 mus musculu
63	37	66.1	510	2	Q6AZW0_BRARE	Q6azw0 brachydanio
64	37	66.1	511	1	ENP1_RAT	P97687 rattus norv
65	37	66.1	539	2	Q8CDV7_MOUSE	Q8cdv7 mus musculu
66	37	66.1	579	2	Q65J87_BACLD	Q65j87 bacillus li
67	37	66.1	605	2	Q8NQY9_CORGL	Q8nqy9 corynebacte
68	37	66.1	621	2	Q82KB5_STRAW	Q82kb5 streptomyce
69	37	66.1	627	1	UVRC_ANASP	Q8yti9 anabaena sp
70	37	66.1	627	2	Q5LPF8_SILPO	Q5lpf8 silicibacte
71	37	66.1	667	2	Q51TU2_MAGGR	Q51tu2 magnaporthe
72	37	66.1	679	2	Q5SMQ4_ORYSA	Q5smq4 oryza sativ
73	37	66.1	747	1	ATCS_SYNP7	P37279 synechococc
74	37	66.1	747	2	Q8KPV5_SYNP7	Q8kpv5 synechococc
75	37	66.1	747	2	Q5N650_SYNP6	Q5n650 synechococc
76	37	66.1	753	2	Q8ZS77_ANASP	Q8zs77 anabaena sp
77	37	66.1	753	2	Q8YWI6_ANASP	Q8ywi6 anabaena sp
78	37	66.1	814	2	Q4FRC3_9GAMM	Q4frc3 psychrobact
79	37	66.1	896	2	Q8TMR9_METAC	Q8tmr9 methanosarc
80	37	66.1	907	2	Q26312_STRPU	Q26312 strongyloce
81	37	66.1	948	2	Q8KQL9_SACER	Q8kql9 saccharopol
82	37	66.1	973	2	Q6N989_RHOPA	Q6n989 rhodopseude
83	37	66.1	1067	1	TRI1_STRCO	Q9rde2 streptomyce
84	37	66.1	1258	2	Q6GUE5_9GAMM	Q6gue5 halomonas m
85	37	66.1	1752	2	Q07265_STRPU	Q07265 strongyloce
86	37	66.1	2936	2	Q9NKP7_LEIMA	Q9nkp7 leishmania
87	36	64.3	67	2	Q9PED6_XYLFA	Q9ped6 xylella fas
88	36	64.3	96	2	Q7TM46_9ZZZZ	Q7tm46 uncultured
89	36	64.3	115	2	Q61XC2_CABER	Q61xc2 caenorhabdi
90	36	64.3	141	2	Q700Q3_PSEPU	Q700q3 pseudomonas
91	36	64.3	160	2	P74976_XYLFA	P74976 xylella fas
92	36	64.3	166	2	Q5T4D2_HUMAN	Q5t4d2 homo sapien
93	36	64.3	194	1	LYG_EPICO	Q90x99 epinephelus
94	36	64.3	217	2	Q7W3Z4_BORPA	Q7w3z4 bordetella
95	36	64.3	238	2	Q9KYB0_STRCO	Q9kyb0 streptomyce
96	36	64.3	248	2	Q7VUW1_BORPE	Q7vuw1 bordetella
97	36	64.3	253	2	Q7P500_FUSNV	Q7p500 fusobacteri
98	36	64.3	255	2	Q8RH63_FUSNN	Q8rh63 fusobacteri
99	36	64.3	263	1	RPP_XYLFT	Q7wfc6 bordetella
100	36	64.3	263	2	Q4FGL4_XYLFA	Q83wl3 xylella fas
101	36	64.3	263	2	Q96EZ5_HUMAN	Q4fgl4 xylella fas
102	36	64.3	275	2	Q5NDF6_HYDEC	Q96ez5 homo sapien
103	36	64.3	282	2	Q4FGL2_XYLFA	Q5ndf6 hydractinia
104	36	64.3	282	2		Q4fgl2 xylella fas

105 36 64.3 283 1 RPP_XYLFA Q9phk6 xylella fas
106 36 64.3 283 2 Q61X02 CAEBR Q61x02 caenorhabdi
107 36 64.3 283 2 Q4ZZA7_PSESY Q4zza7 pseudomonas
108 36 64.3 318 2 Q4UNX9_XANCP Q4unx9 xanthomonas
109 36 64.3 318 2 Q8P3F7_XANCP Q8p3f7 xanthomonas
110 36 64.3 322 2 Q7R524_GIALA Q7r524 giardia lam
111 36 64.3 326 2 Q4K9E6_PSEF5 Q4k9e6 pseudomonas
112 36 64.3 336 2 Q9RDM2_STRCO Q9rdm2 streptomyce
113 36 64.3 336 2 Q82BZ6_STRAW Q82bz6 streptomyce
114 36 64.3 338 2 Q9RTH5_DEIRA Q9rth5 deinococcu
115 36 64.3 355 2 Q882A5_PSESM Q882a5 pseudomonas
116 36 64.3 356 2 Q9I2R6_PSEAE Q9i2r6 pseudomonas
117 36 64.3 365 2 Q4NVY9_9DELT Q4nv9 anaeromyxob
118 36 64.3 375 2 Q61BM4_CAEBR Q61bm4 caenorhabdi
119 36 64.3 390 2 Q7PUJ5_ANOGA Q7puj5 anopheles g
120 36 64.3 391 2 Q8WR22_ANOGA Q8wr22 anopheles g
121 36 64.3 415 2 Q7TZQ1_MYCBO Q7tzq1 mycobacteri
122 36 64.3 415 2 P71981_MYCTU P71981 mycobacteri
123 36 64.3 424 2 Q55X61_CRYNE Q55x61 cryptococcu
124 36 64.3 424 2 Q5KMS1_CRYNE Q5kms1 cryptococcu
125 36 64.3 436 2 Q5F2N5_RAT Q5f2n5 rattus norv
126 36 64.3 462 2 Q94BM9_ARATH Q94bm9 arabidopsis
127 36 64.3 476 2 Q7SD10_NEUCR Q7sd10 neurospora
128 36 64.3 479 2 Q9LML6_ARATH Q9lml6 arabidopsis
129 36 64.3 517 2 Q8N107_HUMAN Q8n107 homo sapien
130 36 64.3 536 2 Q93JK8_STRCO Q93jk8 streptomyce
131 36 64.3 540 2 Q6C246_YARLI Q6c246 yarrowia li
132 36 64.3 560 2 Q9P226_HUMAN Q9p226 homo sapien
133 36 64.3 593 2 Q4KMP2_HUMAN Q4kmp2 homo sapien
134 36 64.3 595 2 Q5T4B2_HUMAN Q5t4b2 homo sapien
135 36 64.3 629 2 Q4IQB8_GIBZE Q4iqb8 gibberella
136 36 64.3 663 2 Q7SCL7_NEUCR Q7scl7 neurospora
137 36 64.3 673 2 Q6VT14_9VIRU Q6vt14 vibrio para
138 36 64.3 675 2 Q6VT78_9VIRU Q6vt78 vibrio para
139 36 64.3 688 2 Q4I141_GIBZE Q4i141 gibberella
140 36 64.3 694 2 Q4NVS9_9DELT Q4nvs9 anaeromyxob
141 36 64.3 700 1 ACOX3_YARLI Q74936 yarrowia li
142 36 64.3 701 2 Q9U233_CABEL Q9u233 caenorhabdi
143 36 64.3 703 1 Y4II_RHISN P55492 rhizobium s
144 36 64.3 742 2 Q5YVJ8_NOCFA Q5yv8 nocardia fa
145 36 64.3 770 2 Q4IYM9_AZOVI Q4iym9 azotobacter
146 36 64.3 951 2 Q9HTC4_PSEAE Q9htc4 pseudomonas
147 36 64.3 1260 1 LIN1_NYCCO P08548 nycitacebus
148 36 64.3 1648 2 Q4P6D7_USTMA Q4p6d7 ustilago ma
149 35 62.5 44 2 Q4GY73_9TRYP Q4gy73 trypanosoma
150 35 62.5 66 2 Q88GU1_PSEPK Q88gu1 pseudomonas
151 35 62.5 86 2 Q7QWC5_GIALA Q7qwc5 giardia lam
152 35 62.5 91 2 Q5GR82_ALCXX Q5gr82 alcaligenes
153 35 62.5 113 2 Q9KT37_VIBCH Q9kt37 vibrio chol
154 35 62.5 117 2 Q7NXS7_CHRVO Q7nxs7 chromobacte
155 35 62.5 130 2 Q27408_METTH Q27408 methanobact
156 35 62.5 130 2 Q58XA3_HCMV Q58ka3 human cytom
157 35 62.5 130 2 Q8QH4_Q_HCMV Q8qh4 human cytom
158 35 62.5 130 2 Q8QH5_HCMV Q8qh5 human cytom
159 35 62.5 130 2 Q8QR26_HCMV Q8qr26 human cytom
160 35 62.5 130 2 Q6SW19_HCMV Q6sw19 human cytom
161 35 62.5 130 2 Q6SWL0_HCMV Q6swl0 human cytom
162 35 62.5 130 2 Q6SWU3_HCMV Q6swu3 human cytom
163 35 62.5 130 2 Q6SWX1_HCMV Q6swx1 human cytom
164 35 62.5 158 2 Q7VX29_BORPE Q7vx29 bordetella
165 35 62.5 158 2 Q7W7Y5_BORPA Q7w7y5 bordetella
166 35 62.5 158 2 Q7WLC5_BORBR Q7wlc5 bordetella
167 35 62.5 195 2 Q8U3H5_PYRFU Q8u3h5 pyrococcus
168 35 62.5 213 2 Q88JD7_PSEPK Q88jd7 pseudomonas
169 35 62.5 226 2 Q82QP7_STRAP Q82qp7 streptomyce
170 35 62.5 227 2 Q836F0_ENTFA Q836f0 enterococcu
171 35 62.5 228 2 Q5TN19_ANOGA Q5tni9 anopheles g
172 35 62.5 228 2 Q6N7U1_RHOPA Q6n7u1 rhodopseudo
173 35 62.5 233 2 Q5ZA10_ORYSA Q5za10 oryza sativ
174 35 62.5 247 2 Q4NRG0_9DELT Q4nrg0 anaeromyxob
175 35 62.5 251 2 Q61U32_CAEBR Q61u32 caenorhabdi
176 35 62.5 252 2 Q7WQD4_BORBR Q7wqd4 bordetella
177 35 62.5 252 2 Q7WCD0_BORPA Q7wcd0 bordetella

178 35 62.5 252 2 Q7VTS9_BORPE Q7vts9 bordetella
179 35 62.5 257 2 Q8YUW3_ANASP Q8yuw3 anabaena sp
180 35 62.5 261 2 Q6ZG10_ORYSA Q6zgl0 oryza sativ
181 35 62.5 264 2 Q4ZVM0_PSESY Q4zvm0 pseudomonas
182 35 62.5 264 2 Q4IWX7_AZOVI Q4iwx7 azotobacter
183 35 62.5 264 2 Q87YS1_PSESM Q87ysl pseudomonas
184 35 62.5 265 2 Q89V33_BRAJA Q89v33 bradyrhizob
185 35 62.5 273 2 Q4MYU9_THEPA Q4myu9 theileria p
186 35 62.5 276 2 Q6Z378_ORYSA Q6z378 oryza sativ
187 35 62.5 283 2 Q87V57_PSESM Q87v57 pseudomonas
188 35 62.5 290 2 Q67QM5_SYMTH Q67qm5 symbiobacte
189 35 62.5 294 2 Q92NI0_RHIME Q92ni0 rhizobium m
190 35 62.5 300 2 P74651_SYNY3 P74651 synecocyst
191 35 62.5 302 2 Q9CBP4_MYCLE Q9cbp4 mycobacteri
192 35 62.5 309 2 Q5WBZ3_BACSK Q5wbz3 bacillus cl
193 35 62.5 312 2 Q829K6_STRAW Q829k6 streptomyce
194 35 62.5 319 2 Q8MVC6_IXOSC Q8mvc6 ixodes scap
195 35 62.5 332 2 Q69231_BACI Q69231 bacillus sp
196 35 62.5 335 2 Q4JYV1_CORJK Q4jyv1 corynebacte
197 35 62.5 341 2 Q9ZET3_XANP2 Q9zet3 xanthobacte
198 35 62.5 354 2 Q5LNT0_SILPO Q5lnt0 silicibacte
199 35 62.5 360 2 Q7URK3_RHOBA Q7urk3 rhodopirell
200 35 62.5 361 2 Q6LHL8_PHOPR Q6lhl8 photobacter
201 35 62.5 364 1 HEMH_CHRVO Q7nv65 chromobacte
202 35 62.5 370 2 Q9HYJ0_PSEAE Q9hyj0 pseudomonas
203 35 62.5 379 2 Q4J471_AZOVI Q4j471 azotobacter
204 35 62.5 383 2 Q6I609_ORYSA Q6i609 oryza sativ
205 35 62.5 387 2 Q98FX1_RHILO Q98fx1 rhizobium l
206 35 62.5 396 2 Q943M0_ORYSA Q943m0 oryza sativ
207 35 62.5 405 2 Q6RK21_ERWCT Q6rk21 erwinia car
208 35 62.5 412 2 Q9VZV5_DROME Q9vzv5 drosophila
209 35 62.5 426 2 Q4ZRY2_PSESY Q4zry2 pseudomonas
210 35 62.5 426 2 Q880G9_PSESM Q880g9 pseudomonas
211 35 62.5 427 1 CYB6_CHLLT Q59297 chlorobium
212 35 62.5 429 2 Q6Z370_ORYSA Q6z370 oryza sativ
213 35 62.5 437 2 Q765A8_9PEZI Q765a8 coniochaeta
214 35 62.5 468 2 Q4I330_GIBZE Q4i330 gibberella
215 35 62.5 480 2 Q8KF93_CHLTE Q8kf93 chlorobium
216 35 62.5 482 2 O04201_ARATH O04201 arabidopsis
217 35 62.5 495 2 Q9H9A9_HUMAN Q9h9a9 homo sapien
218 35 62.5 495 2 Q84CJ1_9ACTO Q84cj1 streptomyce
219 35 62.5 520 2 Q9H0W8_HOMAN Q9h0w8 homo sapien
220 35 62.5 520 2 Q988D6_RHILO Q988d6 rhizobium l
221 35 62.5 520 2 Q5PQS6_RAT Q5pq6 rattus norv
222 35 62.5 520 2 Q8BYJ3_MOUSE Q8byj3 mus musculu
223 35 62.5 520 2 Q9DB90_MOUSE Q9db90 mus musculu
224 35 62.5 527 2 Q4ZTC1_PSESY Q4ztcl pseudomonas
225 35 62.5 534 2 Q4SVX6_TETNG Q4svx6 tetraodon n
226 35 62.5 542 2 Q55KX2_CRYNE Q55kx2 cryptococcu
227 35 62.5 542 2 Q5KAK7_CRYNE Q5kak7 cryptococcu
228 35 62.5 545 2 Q6CAR7_YARLI Q6car7 yarrowia li
229 35 62.5 547 2 Q55KX1_CRYNE Q55kx1 cryptococcu
230 35 62.5 547 2 Q5KAK8_CRYNE Q5kak8 cryptococcu
231 35 62.5 553 2 Q55YK8_CRYNE Q55yk8 cryptococcu
232 35 62.5 553 2 Q5KLD0_CRYNE Q5kld0 cryptococcu
233 35 62.5 557 2 Q55NQ3_CRYNE Q55nq3 cryptococcu
234 35 62.5 557 2 Q5KC38_CRYNE Q5kc38 cryptococcu
235 35 62.5 586 2 Q4LXI4_9BURK Q4lxi4 burkholderi
236 35 62.5 607 1 GLTL4_HUMAN Q6p9a2 h putative
237 35 62.5 607 2 Q58A54_HUMAN Q58a54 homo sapien
238 35 62.5 607 2 Q59J91_MOUSE Q59j91 mus musculu
239 35 62.5 622 1 GLTL4_MOUSE Q8klb9 m putative
240 35 62.5 641 2 Q63JQ4_BURPS Q63jq4 burkholderi
241 35 62.5 641 2 Q62AN3_BURMA Q62an3 burkholderi
242 35 62.5 647 2 Q82L15_STRAW Q82l15 streptomyce
243 35 62.5 697 1 EFG2_SHEON Q8ei7 shewanella
244 35 62.5 702 2 Q52ET5_MAGGR Q52et5 magnaporthe
245 35 62.5 702 2 Q7UY61_RHOBA Q7uy61 rhodopirell
246 35 62.5 723 2 Q51ND2_MAGGR Q51nd2 magnaporthe
247 35 62.5 723 2 Q4UXI0_XANCP Q4uxi0 xanthomonas
248 35 62.5 723 2 Q8P6N1_XANCP Q8p6n1 xanthomonas
249 35 62.5 727 2 Q8PHY5_XANAC Q8phy5 xanthomonas
250 35 62.5 731 2 Q6N6R0_RHOPA Q6n6r0 rhodopseudo

397	34	60.7	618	2	Q8W4C3	ARATH	Q8w4c3	arabidopsis
398	34	60.7	644	2	Q7XND4	ORYSA	Q7xnd4	oryza sativ
399	34	60.7	647	2	Q825E9	STRAW	Q825e9	streptomyce
400	34	60.7	659	2	Q5ZCG0	ORYSA	Q5zcg0	oryza sativ
401	34	60.7	660	2	Q17248	BOOMI	Q17248	boophilus m
402	34	60.7	693	2	Q8ZQ94	SALTY	Q8zq94	salmonella
403	34	60.7	703	2	Q4WCY1	ASPFU	Q4wcy1	aspergillus
404	34	60.7	725	2	Q5XVM9	ENTFA	Q5xvm9	enterococcu
405	34	60.7	733	2	Q89LV4	BRAJA	Q89lv4	bradyrhizob
406	34	60.7	738	2	Q8XYC9	RALSO	Q8xyc9	ralstonia s
407	34	60.7	750	2	Q9FGC2	ARATH	Q9fgc2	arabidopsis
408	34	60.7	856	2	Q4Q466	LEIMA	Q4q466	leishmania
409	34	60.7	859	2	Q9FW70	ORYSA	Q9fw70	oryza sativ
410	34	60.7	870	2	Q7QJD3	ANOGA	Q7qjd3	anopheles g
411	34	60.7	901	1	MALT_SALT1		Q8z227	salmonella
412	34	60.7	901	1	MALT_SALTY		Q8zli3	salmonella
413	34	60.7	901	2	Q57IV9	SALCH	Q57iv9	salmonella
414	34	60.7	901	2	Q5PM00	SALPA	Q5pm00	salmonella
415	34	60.7	905	2	Q8XRT2	RALSO	Q8xrt2	ralstonia s
416	34	60.7	906	2	Q630Y0	BACCZ	Q630y0	bacillus ce
417	34	60.7	987	2	Q7YZM8	CABEL	Q7yzm8	caenorhabdi
418	34	60.7	1051	2	Q9Y150	DROME	Q9y150	drosophila
419	34	60.7	1058	2	Q4WV56	ASPFU	Q4wv56	aspergillus
420	34	60.7	1076	2	Q605X2	METCA	Q605x2	methylococc
421	34	60.7	1098	2	Q7Z146	CABEL	Q7z146	caenorhabdi
422	34	60.7	1108	2	Q4X0Z7	ASPFU	Q4x0z7	aspergillus
423	34	60.7	1112	2	Q8GD29	BORAV	Q8gd29	bordetella
424	34	60.7	1115	2	Q9LSV7	ARATH	Q9lsv7	arabidopsis
425	34	60.7	1116	2	Q53VZ7	THET8	Q53vz7	thermus the
426	34	60.7	1159	1	RGYR2_AQUAE		Q67226	aquifex ao
427	34	60.7	1166	1	RGYR2_SULSO		Q97zf5	sulfolobus
428	34	60.7	1166	1	RGYR_SULSH		P74759	sulfolobus
429	34	60.7	1192	2	Q610E4	CAEBR	Q610e4	caenorhabdi
430	34	60.7	1208	2	Q7RWM1	NEUCR	Q7rwm1	neurospora
431	34	60.7	1214	1	RGYR_PYRAB		Q9uz86	pyrococcus
432	34	60.7	1214	1	RGYR_PYRFU		P95479	pyrococcus
433	34	60.7	1221	1	Q49600	METKA	Q49600	methanopyru
434	34	60.7	1222	1	RGYR2_AERPE		Q9yc75	aeropyrum p
435	34	60.7	1228	1	RGYR_PYRAE		Q8zxt5	pyrobaculum
436	34	60.7	1332	1	YSY1_CABEL		Q20255	caenorhabdi
437	34	60.7	1379	2	Q581S3	9ACTO	Q58is3	streptomyce
438	34	60.7	1529	2	Q9V7C6	ASPTPE	Q9y7c6	aspergillus
439	34	60.7	1613	1	RGYR_METJA		Q58907	methanococc
440	34	60.7	1624	1	RGYR_PYRHO		Q58530	pyrococcus
441	34	60.7	1711	1	RGYR_PYRKO		Q6f598	pyrococcus
442	34	60.7	1897	2	Q5L2K6	GEOKA	Q5l2k6	geobacillus
443	34	60.7	2066	2	Q9Z3T8	PSESX	Q9z3t8	pseudomonas
444	34	60.7	2066	2	Q87W69	PSESM	Q87w69	pseudomonas
445	34	60.7	2090	2	Q9P6X4	NEUCR	Q9p6x4	neurospora
446	34	60.7	2219	2	Q4P405	USTMA	Q4p405	ustilago ma
447	34	60.7	4437	2	Q8GGP3	STRAZ	Q8ggp3	streptomyce
448	33.5	59.8	152	2	Q4KPK2	9PSED	Q4kpk2	pseudomonas
449	33.5	59.8	209	2	Q65HI1	BACLD	Q65hi1	bacillus li
450	33.5	59.8	418	2	Q9ERI0	9RODE	Q9eri0	microtus br
451	33.5	59.8	640	2	Q5TVL2	ANOGA	Q5tv12	anopheles g
452	33.5	59.8	688	2	Q6V7N1	9CAUD	Q6v7n1	burkholderi
453	33	58.9	38	2	Q8KCI1	CHLTE	Q8kc11	chlorobium
454	33	58.9	68	2	Q8UKE3	AGRT5	Q8uke3	agrobacteri
455	33	58.9	90	2	Q4IU87	AZOVI	Q4iu87	azotobacter
456	33	58.9	104	2	Q72CX5	DESVH	Q72cx5	desulfovibr
457	33	58.9	112	2	Q6N0E5	RHOPA	Q6n0e5	rhodopseudo
458	33	58.9	112	2	Q4R9Q6	TETNG	Q4r9q6	tetraodon n
459	33	58.9	114	2	Q6W5R6	TRYCR	Q6w5r6	trypanosoma
460	33	58.9	114	2	Q7W775	BORPA	Q7w775	bordetella
461	33	58.9	114	2	Q7WIJ7	BORBR	Q7wij7	bordetella
462	33	58.9	125	2	Q5KFE3	CRYNE	Q5kfe3	cryptococcu
463	33	58.9	130	2	Q8VMF1	PSEPU	Q8vmf1	pseudomonas
464	33	58.9	140	2	Q7P0I8	CHRVO	Q7p0i8	chromobacte
465	33	58.9	142	2	Q6N7J4	RHOPA	Q6n7j4	rhodopsendo
466	33	58.9	146	2	Q9Y9H4	AERPE	Q9y9h4	aeropyrum p
467	33	58.9	148	2	Q6IIP0	DROME	Q6iip0	drosophila
468	33	58.9	149	2	Q55QSI	CRYNE	Q55qsl	cryptococcu
469	33	58.9	151	2	Q67MR0	SYNTH	Q67mr0	symbiobacte

470	33	58.9	154	2	Q8TLD3	METAC	Q8tld3	methanosarc
471	33	58.9	154	2	Q4LYN3	9BURK	Q4lyn3	burkholderi
472	33	58.9	179	2	Q9HZH7	PSEAE	Q9hzh7	pseudomonas
473	33	58.9	181	2	Q4RZR7	TETNG	Q4rzt7	tetraodon n
474	33	58.9	185	2	Q5JDD5	PYRKO	Q5jdd5	pyrococcus
475	33	58.9	187	2	Q629E8	CAEBR	Q629e8	caenorhabdi
476	33	58.9	190	2	Q9UYS6	PYRAB	Q9uys6	pyrococcus
477	33	58.9	196	2	Q7WZ84	9ACTO	Q7wz84	nonomuraea
478	33	58.9	196	2	Q5X2K8	LEGPA	Q5x2k8	legionella
479	33	58.9	197	1	HAM1_CHRVO		Q7nzj6	chromobacte
480	33	58.9	199	2	Q9AUC9	SESIN	Q9auc9	sesamum ind
481	33	58.9	200	2	Q51RL6	MAGGR	Q51rl6	magnaporthes
482	33	58.9	200	2	Q7FAY5	ORYSA	Q7fay5	oryza sativ
483	33	58.9	200	2	Q828G1	STRAW	Q828g1	streptomyce
484	33	58.9	200	2	Q8CJS4	STRCO	Q8cjs4	streptomyce
485	33	58.9	214	2	Q7VT72	BORPE	Q7vt72	bordetella
486	33	58.9	215	2	Q96Y57	SULTO	Q96y57	sulfolobus
487	33	58.9	218	2	Q4WH75	ASPFU	Q4wh75	aspergillus
488	33	58.9	224	2	Q51KN9	MAGGR	Q51kn9	magnaporthes
489	33	58.9	225	2	Q9HVR1	PSEAE	Q9hvr1	pseudomonas
490	33	58.9	232	2	Q840V0	SPHEL	Q840v0	sphingomona
491	33	58.9	240	2	Q57891	PYRHO	Q57891	pyrococcus
492	33	58.9	240	2	Q6CL65	KLULA	Q6cl65	kluveromyc
493	33	58.9	243	2	Q7NN94	GLOVI	Q7nn94	gloeobacter
494	33	58.9	244	2	Q9RWM1	DEIRA	Q9rwm1	deinococcus
495	33	58.9	248	2	Q5P7A3	AZOSE	Q5p7a3	azoarcus sp
496	33	58.9	249	2	Q8X6D3	ECO57	Q8x6d3	escherichia
497	33	58.9	257	2	Q5LU90	SILPO	Q5lu90	silicibacte
498	33	58.9	258	2	Q4PMJ6	IXOSC	Q4pmj6	ixodes scap
499	33	58.9	261	2	Q87AQ1	XYLFT	Q87aq1	xyella fas
500	33	58.9	261	2	Q9PEW5	XYLFA	Q9pew5	xyella fas
501	33	58.9	261	2	Q7X4S4	BACLI	Q7x4s4	bacillus li
502	33	58.9	261	2	Q65FM6	BACLD	Q65fm6	bacillus li
503	33	58.9	265	2	Q8KJ89	RHILO	Q8kj89	rhizobium l
504	33	58.9	272	2	Q7TW42	MYCBO	Q7tw42	mycobacteri
505	33	58.9	272	2	O06282	MYCTU	O06282	mycobacteri
506	33	58.9	273	2	Q5P6P9	AZOSE	Q5p6p9	azoarcus sp
507	33	58.9	275	2	Q9A3U3	CAUCR	Q9a3u3	caulobacter
508	33	58.9	281	2	Q63X73	BURPS	Q63x73	burkholderi
509	33	58.9	282	1	FPG_PROMA		Q7vdk6	prochloroco
510	33	58.9	287	2	Q7TXN3	MYCBO	Q7txn3	mycobacteri
511	33	58.9	287	2	Q7D6F5	MYCTU	Q7d6f5	mycobacteri
512	33	58.9	288	2	Q5PBF3	ANAMM	Q5pbf3	anaplasma m
513	33	58.9	288	2	Q8Y216	RALSO	Q8y216	ralstonia s
514	33	58.9	290	2	Q5YVI9	NOCFA	Q5yvi9	nocardia fa
515	33	58.9	292	2	Q62ML9	BURMA	Q62ml9	burkholderi
516	33	58.9	293	2	Q8PII4	XANAC	Q8pii4	xanthomonas
517	33	58.9	295	2	Q4LU32	9BURK	Q4lu32	burkholderi
518	33	58.9	299	2	Q4TMS3	9SPHN	Q4tms3	erythrobacte
519	33	58.9	302	2	Q7NV73	CHRVO	Q7nv73	chromobacte
520	33	58.9	305	1	RNZ_METMA		Q8q032	methanosarc
521	33	58.9	310	2	Q5NNL6	ZYMMO	Q5nnl6	zymomonas m
522	33	58.9	315	2	Q69LV5	ORYSA	Q69lv5	oryza sativ
523	33	58.9	316	2	O34689	BACSU	O34689	bacillus su
524	33	58.9	318	2	Q742P0	MYCPA	Q742p0	mycobacteri
525	33	58.9	320	2	Q4PMM5	IXOSC	Q4pmm5	ixodes scap
526	33	58.9	327	2	Q4WCE5	ASPFU	Q4wce5	aspergillus
527	33	58.9	327	2	Q4ILG2	GIBZE	Q4ilg2	gibberella
528	33	58.9	330	2	Q51LI0	MAGGR	Q51li0	magnaporthes
529	33	58.9	330	2	Q4TQX1	9SPHN	Q4tqx1	erythrobact
530	33	58.9	331	2	Q4R3B8	MACFA	Q4r3b8	macaca faasc
531	33	58.9	333	2	Q9RJ10	STRCK	Q9rj10	streptomyce
532	33	58.9	333	2	Q88LW1	PSEPK	Q88lw1	pseudomonas
533	33	58.9	335	2	Q7SHA6	NEUCR	Q7sha6	neurospora
534	33	58.9	335	2	Q870E8	9EURO	Q870e8	penicillium
535	33	58.9	335	2	Q9ZWF3	LYCES	Q9zwf3	lycoperisico
536	33	58.9	342	2	Q9X6D4	CLOBE	Q9x6d4	clostridium
537	33	58.9	354	2	Q8PZ82	METMA	Q8pz82	methanosarc
538	33	58.9	355	2	Q4ZTM6	PSESY	Q4ztm6	pseudomonas
539	33	58.9	357	2	Q6LXY5	METMP	Q6lxy5	methanococc
540	33	58.9	357	2	Q9A4W7	CAUCR	Q9a4w7	caulobacter
541	33	58.9	359	2	Q852B6	ORYSA	Q852b6	oryza sativ
542	33	58.9	360	2	Q9HR47	HALSA	Q9hr47	halobacteri

Q8tld3	methanosarc
Q4lyn3	burkholderi
Q9hzh7	pseudomonas
Q4rzt7	tetraodon n
Q5jdd5	pyrococcus
Q629e8	caenorhabdi
Q9uys6	pyrococcus
Q7wz84	nonomuraea
Q5x2k8	legionella
Q7nzj6	chromobacte
Q9auc9	sesamum ind
Q51rl6	magnaporthes
Q7fay5	oryza sativ
Q828g1	streptomyce
Q8cjs4	streptomyce
Q7vt72	bordetella
Q96y57	sulfolobus
Q4wh75	aspergillus
Q51kn9	magnaporthes
Q9hvr1	pseudomonas
Q840v0	sphingomona
Q57891	pyrococcus
Q6cl65	kluveromyc
Q7nn94	gloeobacter
Q9rwm1	deinococcus
Q5p7a3	azoarcus sp
Q8x6d3	escherichia
Q5lu90	silicibacte
Q4pmj6	ixodes scap
Q87aq1	xylella fas
Q9ypew5	xylella fas
Q7x4s4	bacillus li
Q65fm6	bacillus li
Q8kj89	rhizobium l
Q7tw42	mycobacteri
Q06282	mycobacteri
Q5p6p9	azoarcus sp
Q9a3u3	caulobacter
Q63x73	burkholderi
Q7vdk6	prochloroco
Q7txn3	mycobacteri
Q7d6f5	mycobacteri
Q5pbf3	anaplasma m
Q8y216	racistonia s
Q5Yv19	nocardia fa
Q62ml9	burkholderi
Q8pi14	xanthomonas
Q4lu32	burkholderi
Q4tms3	erythrobact
Q7nv73	chromobacte
Q8q032	methanosarc
Q5nnl6	zymomonas m
Q69lv5	oryza sativ
Q34689	bacillus su
Q742p0	mycobacteri
Q4pmm5	ixodes scap
Q4wces	aspergillus
Q4ilg2	gibberella
Q51li0	magnaporthes
Q4tqx1	erythrobact
Q4r3b8	macaca fasc
Q9rj10	streptomyces
Q88l1w	pseudomonas
Q7sha6	neurospora
Q870e8	penicillium
Q9zwf3	lycopersidum
Q9x6d4	clostridium
Q8pz82	methanosarc
Q4ztm6	pseudomonas
Q6lxys	methanococc
Q9a4w7	caulobacter
Q85a67	oryzasativ
Q9hr47	halobacteri

543	33	58.9	362	2	Q6J657_ORYSA	Q6j657_oryza sativ	616	33	58.9	570	2	Q7S2U7_NEUCR	Q7s2u7_neurospora
544	33	58.9	362	2	Q5Y0T7_9ALPH	Q5y0t7_cercopithec	617	33	58.9	573	1	MLO11_ARATH	Q9fi00_arabidopsis
545	33	58.9	363	2	Q8NQB0_CORGL	Q8nqb0_corynebacte	618	33	58.9	587	2	Q4W8R1_WHEAT	Q4w8r1_triticum ae
546	33	58.9	364	2	Q9R701_9RHIZ	Q9r701_agrobacteri	619	33	58.9	589	2	Q4W8R0_WHEAT	Q4w8r0_triticum ae
547	33	58.9	364	2	Q63PI5_BURPS	Q63pi5_burkholderi	620	33	58.9	601	2	Q4H6I3_9DEIO	Q4h6i3_deinococcu
548	33	58.9	365	2	Q4P6I7_USTMA	Q4p6i7_ustilago ma	621	33	58.9	603	2	Q8FPW3_COREF	Q8fpw3_corynebacte
549	33	58.9	367	2	Q74GD2_GEOSL	Q74gd2_geobacter s	622	33	58.9	603	2	Q9L2I7_STRCO	Q9l2i7_streptomyce
550	33	58.9	367	2	Q98IX0_RHILO	Q98ix0_rhizobium l	623	33	58.9	604	2	Q5B4E5_EMENI	Q5b4e5_aspergillus
551	33	58.9	371	2	Q6YXI7_ORYSA	Q6yxi7_oryza sativ	624	33	58.9	605	2	Q4NTT6_9DELT	Q4ntt6_aeromonas
552	33	58.9	376	2	Q6AFB4_LEIXX	Q6afe4_leifsonia x	625	33	58.9	608	2	Q55WN2_CRYNE	Q55wn2_cryptococcu
553	33	58.9	377	2	Q82KPO_STRAW	Q82kp0_streptomyce	626	33	58.9	608	2	Q5KJL3_CRYNE	Q5kjl3_cryptococcu
554	33	58.9	377	2	Q6ASQ2_ORYSA	Q6asq2_oryza sativ	627	33	58.9	610	2	Q6BMT2_DEBHA	Q6bmt2_debaryomyce
555	33	58.9	378	2	Q6ASQ2_ORYSA	Q6asq2_oryza sativ	628	33	58.9	610	2	Q8XLR0_CLOPE	Q8xlr0_clostridium
556	33	58.9	384	1	RECF_BRUME	Q8yed7_brucella me	629	33	58.9	620	2	Q4WIF1_ASPFU	Q4wif1_aspergillus
557	33	58.9	384	1	RECF_BRUSU	Q8g3e5_brucella su	630	33	58.9	626	2	Q6C6T6_YARLI	Q6c6t6_yarrowia li
558	33	58.9	384	2	Q57G08_BRUAB	Q57g08_brucella ab	631	33	58.9	630	2	Q642H1_BRARE	Q642h1_brachydanio
559	33	58.9	389	2	Q7MSI7_WOLSU	Q7msi7_wolinella s	632	33	58.9	632	2	Q4SI64_TETNG	Q4si64_tetrahodon n
560	33	58.9	389	2	Q89NU2_BRAJA	Q89nu2_bradyrhizob	633	33	58.9	637	2	Q4SI64_TETNG	Q4si64_tetrahodon n
561	33	58.9	390	2	Q4S165_TETNG	Q4s165_tetrahodon n	634	33	58.9	639	2	Q4I245_AZOVI	Q4i245_azotobacter
562	33	58.9	391	2	Q60ER6_ORYSA	Q60er6_oryza sativ	635	33	58.9	640	2	Q7RW79_NEUCR	Q7rw79_neurospora
563	33	58.9	396	1	SOTB_HAEIN	P44535_haemophilus	636	33	58.9	642	2	Q5AQE5_EMENI	Q5aqe5_aspergillus
564	33	58.9	396	2	Q9HS71_HALSA	Q9hs71_halobacteri	637	33	58.9	642	2	Q4URL8_XANCP	Q4url8_xanthomonas
565	33	58.9	402	2	Q4NNL7_9DELT	Q4nnl7_aeromonas	638	33	58.9	656	2	Q8PBY3_XANCP	Q8pb33_xanthomonas
566	33	58.9	402	2	Q4QP52_HAB18	Q4qp52_haemophilus	639	33	58.9	656	2	Q8PNI2_XANAC	Q8pni2_xanthomonas
567	33	58.9	402	2	Q5YR35_NOCFA	Q5yr35_nocardia fa	640	33	58.9	665	2	Q52C86_MAGGR	Q52c86_magnaportha
568	33	58.9	404	2	Q5BDA3_EMENI	Q5bda3_aspergillus	641	33	58.9	666	2	Q27798_METTH	Q27798_methanobact
569	33	58.9	405	2	Q62FR1_BURMA	Q62fr1_burkholderi	642	33	58.9	666	2	Q6VPU8_DROVI	Q6vp8_drosophila
570	33	58.9	407	2	Q4HU19_GIBZE	Q4hu19_gibberella	643	33	58.9	668	2	Q8YWX6_ANASP	Q8ywx6_anabaena sp
571	33	58.9	421	2	Q89LW8_BRAJA	Q89lw8_bradyrhizob	644	33	58.9	674	2	Q5SYA4_CRYNE	Q5sya4_cryptococcu
572	33	58.9	422	2	Q6NJM3_CORDI	Q6njm3_corynebacte	645	33	58.9	675	2	Q5KLN4_CRYNE	Q5kln4_cryptococcu
573	33	58.9	428	1	CINA_MYCLE	Q9x7d6_mycobacteri	646	33	58.9	679	1	YHC9_YEAST	Q5h430_xanthomonas
574	33	58.9	434	2	Q5LV46_SILPO	Q5lv46_silicibacte	647	33	58.9	682	2	Q8A5J3_BACTN	P38738_saccharomyc
575	33	58.9	436	2	Q4JU84_CORJK	Q4ju84_corynebacte	648	33	58.9	687	2	Q63MU8_BURPS	Q8a5j3_bacteroides
576	33	58.9	440	2	Q59XE2_CANAL	Q59xe2_candida alb	649	33	58.9	687	2	Q62DI2_BURKH	Q63mu8_burkholderi
577	33	58.9	442	2	Q4H8M9_9DEIO	Q4h8m9_deinococcus	650	33	58.9	688	2	Q7V7K3_PROMM	Q62di2_burkholderi
578	33	58.9	443	2	Q7QJC9_ANOGA	Q7qjc9_anopheles g	651	33	58.9	703	2	Q7U741_SYNPX	Q7v7k3_prochloroco
579	33	58.9	446	2	Q4K6Q8_PSEF5	Q4k6q8_pseudomonas	652	33	58.9	707	1	DCOR_LEIDO	Q7u741_synehococc
580	33	58.9	447	2	Q4P9P8_USTMA	Q4p9p8_ustilago ma	653	33	58.9	707	1	DCOR_LEIDO	P27116_leishmania
581	33	58.9	447	2	Q4TAV8_TETNG	Q4tav8_tetrahodon n	654	33	58.9	707	2	Q4QGR5_LEIMA	P27116_leishmania
582	33	58.9	448	2	Q9LJA6_ARATH	Q9lja6_arabidopsis	655	33	58.9	709	2	Q4PE98_USTMA	Q4qgr5_leishmania
583	33	58.9	450	2	Q6BIB0_DEBHA	Q6bib0_debaryomyce	656	33	58.9	715	2	Q7VP04_HAEDU	Q4pe98_ustilago ma
584	33	58.9	450	2	Q6C655_YARLI	Q6c655_yarrowia li	657	33	58.9	721	2	P90517_CRIFA	Q7vp04_haemophilus
585	33	58.9	451	2	Q522V1_MAGGR	Q522v1_magnaportha	658	33	58.9	728	2	Q6DCW9_XENLA	P90517_crithidia f
586	33	58.9	452	2	Q5FQI0_GLUOX	Q5fq10_gluconobact	659	33	58.9	732	2	Q9Y0C4_LEITA	Q6dcw9_xenopus lae
587	33	58.9	455	2	Q5DL07_9INFA	Q5dl07_influenza a	660	33	58.9	735	2	Q4P0G8_USTMA	Q9y0c4_leishmania
588	33	58.9	455	2	Q5DL08_9INFA	Q5dl08_influenza a	661	33	58.9	738	2	Q7SH46_NEUCR	Q4p0g8_ustilago ma
589	33	58.9	455	2	Q5DL09_9INFA	Q5dl09_influenza a	662	33	58.9	741	2	Q4HC03_9DEIO	Q7sh46_neurospora
590	33	58.9	458	2	Q4LRS8_9BURK	Q4lrs8_burkholderi	663	33	58.9	746	2	Q51KR6_MAGGR	Q4hc03_deinococcus
591	33	58.9	461	2	Q5DL10_9INFA	Q5dl10_influenza a	664	33	58.9	751	1	CNOT3_MOUSE	Q51kr6_magnaportha
592	33	58.9	462	2	Q4PLH8_FUSPR	Q4plh8_fusarium pr	665	33	58.9	753	1	CNOT3_HUMAN	Q8k0v4_mus musculu
593	33	58.9	463	2	Q4I403_GIBZE	Q4i403_gibberella	666	33	58.9	765	2	Q8XZK9_RALSO	O75175_homo sapien
594	33	58.9	466	2	Q6Y328_TRYCR	Q6y328_trypanosoma	667	33	58.9	774	2	Q9X6W0_PSEAE	Q8xzk9_ralstonia s
595	33	58.9	469	2	Q5NA57_ORYSA	Q5na57_oryza sativ	668	33	58.9	775	2	Q4P2I1_USTMA	Q9x6w0_pseudomonas
596	33	58.9	469	2	Q6GYV2_9INFA	Q6gyv2_influenza a	669	33	58.9	779	2	Q94DS6_ORYSA	Q4p2i1_ustilago ma
597	33	58.9	469	2	Q6XV53_9INFA	Q6xv53_influenza a	670	33	58.9	780	2	O50054_SPIOL	Q94ds6_oryza sativ
598	33	58.9	472	2	Q6LM64_PHOPR	Q6lm64_photobacter	671	33	58.9	815	2	Q6MXP8_SERMA	O50054_spinacia ol
599	33	58.9	474	2	Q6CX40_KLULA	Q6cx40_kluyveromyc	672	33	58.9	815	2	Q6U5P1_KLEPN	Q6mxp8_serratia ma
600	33	58.9	474	2	Q6CX40_KLULA	Q6cx40_kluyveromyc	673	33	58.9	816	2	Q5JMF3_ORYSA	Q6u5p1_klebsiella
601	33	58.9	474	2	Q4RSP5_TETNG	Q4rsp5_tetrahodon n	674	33	58.9	822	2	Q5JMF3_ORYSA	Q5jmf3_oryza sativ
602	33	58.9	476	2	Q87VQ8_PSEMG	Q87vq8_pseudomonas	675	33	58.9	822	2	Q66B31_YERPS	Q5jmf3_orysinia pe
603	33	58.9	478	2	Q70PA5_9DELT	Q70pa5_melittangiu	676	33	58.9	824	1	SILP_SALTY	Q66b31_yersinia ps
604	33	58.9	479	2	Q841L7_9ACTO	Q841l7_streptomyce	677	33	58.9	825	2	Q54X49_DICDI	Q9zhc7_salmonella
605	33	58.9	481	2	Q75BT3_ASHGO	Q75bt3_ashbya goss	678	33	58.9	825	2	Q6EMD7_ECOLI	Q54x49_dictyosteli
606	33	58.9	486	2	Q4NRJ2_9DELT	Q4nrj2_aeromonas	679	33	58.9	843	2	Q8D0D4_YERPE	Q6emd7_escherichia
607	33	58.9	519	2	Q8RSL8_9BACT	Q8rs18_uncultured	680	33	58.9	843	2	Q74UL6_YERPE	Q8d0d4_yersinia pe
608	33	58.9	530	2	Q6CS52_KLULA	Q6cs52_kluyveromyc	681	33	58.9	847	2	Q6TNN2_BRARE	Q74ul6_yersinia pe
609	33	58.9	530	2	Q8NQY7_CORGL	Q8nqy7_corynebacte	682	33	58.9	850	1	PBPA_ECOLI	Q6tnv2_brachydanio
610	33	58.9	532	2	Q5YRX1_NOCFA	Q5yrx1_nocardia fa	683	33	58.9	850	2	Q83PW8_SHIFL	P02918_escherichia
611	33	58.9	538	2	Q6BJS5_DEBHA	Q6bjs5_debaryomyce	684	33	58.9	851	2	Q6CZQ2_ERWCT	Q83pw8_shigella fl
612	33	58.9	538	2	Q4N9C0_THEPA	Q4n9c0_theileria p	685	33	58.9	858	2	Q57IY1_SALCH	Q6czq2_erwinia car
613	33	58.9	541	2	Q50491_STRCO	Q50491_streptomyce	686	33	58.9	858	2	Q5PLY5_SALPA	Q57iy1_salmonella
614	33	58.9	547	1	MUTL_DEIRA	Q9rtr0_deinococcus	687	33	58.9	858	2	Q8ZLJ7_SALTY	Q5ply5_salmonella
615	33	58.9	554	1	MLO14_ARATH	Q94kb1_arabidopsis	688	33	58.9	858	2	Q8ZLJ7_SALTY	Q8z2l1_salmonella
													Q8zlj7_salmonella

689	33	58.9	858	2	Q8FCV2	ECOL6	Q8fcv2	escherichia
690	33	58.9	858	2	Q6MQN7	BDEBA	Q6mqn7	bdellovibri
691	33	58.9	858	2	Q8X809	ECO57	Q8x809	escherichia
692	33	58.9	866	2	Q55AA8	DICDI	Q55aa8	dictyosteli
693	33	58.9	868	2	Q7X7Q3	ORYSA	Q7x7q3	oryza sativ
694	33	58.9	868	2	Q8A2H7	BACTN	Q8a2h7	bacteroides
695	33	58.9	873	2	Q4WI49	ASPFU	Q4wi49	aspergillus
696	33	58.9	875	2	Q7S029	NEUCR	Q7s029	neurospora
697	33	58.9	881	2	Q4ZP54	PSESY	Q4zp54	pseudomonas
698	33	58.9	886	2	Q87X45	PSESM	Q87x45	pseudomonas
699	33	58.9	896	2	Q7UYC9	RHOBA	Q7uyc9	rhodopirell
700	33	58.9	901	1	MALT	ECO57	Q8x701	escherichia
701	33	58.9	901	1	MALT	ECOL6	Q8cxx5	escherichia
702	33	58.9	901	1	MALT	ECOL1	P06993	escherichia
703	33	58.9	902	2	Q6LHM8	PHOPR	Q6lhm8	photobacter
704	33	58.9	906	2	Q9TTJ4	RABIT	Q9ttj4	oryctolagus
705	33	58.9	908	2	Q7ZVL4	BRARE	Q7zvl4	brachydanio
706	33	58.9	933	2	Q4SZM3	TETNG	Q4szm3	tetraodon n
707	33	58.9	934	2	Q4SN87	TETNG	Q4sn87	tetraodon n
708	33	58.9	939	2	Q4ZP59	PSESY	Q4zp59	pseudomonas
709	33	58.9	946	2	Q4NZU4	9DELT	Q4nzu4	anaeromyxob
710	33	58.9	952	2	Q8NIT8	NEUCR	Q8nit8	neurospora
711	33	58.9	973	2	Q6ZR04	HUMAN	Q6zr04	homo sapien
712	33	58.9	982	2	Q7ZYZ1	BRARE	Q7zyz1	brachydanio
713	33	58.9	987	2	Q4SYG9	TETNG	Q4syg9	tetraodon n
714	33	58.9	1007	2	Q8C8T6	MOUSE	Q8c8t6	mus musculu
715	33	58.9	1015	2	Q86JP2	DICDI	Q86jp2	dictyosteli
716	33	58.9	1087	2	Q4HWP8	GIBZE	Q4hwp8	gibberella
717	33	58.9	1118	1	UBP8	HUMAN	P40818	homo sapien
718	33	58.9	1124	2	Q4Q685	LEIMA	Q4q685	leishmania
719	33	58.9	1182	2	Q9Y7C1	MAGGR	Q9y7c1	magnaporthe
720	33	58.9	1182	2	Q51TA9	MAGGR	Q51ta9	magnaporthe
721	33	58.9	1211	2	Q76YV0	9CAUD	Q76yv0	bacterioph
722	33	58.9	1365	2	Q5AZ14	EMENI	Q5az14	aspergillus
723	33	58.9	1426	2	Q522R0	MAGGR	Q522r0	magnaporthe
724	33	58.9	1468	2	Q7XH59	ORYSA	Q7xh59	oryza sativ
725	33	58.9	1468	2	Q8W5K9	ORYSA	Q8w5k9	oryza sativ
726	33	58.9	1483	2	Q5S864	CRYNE	Q5s864	cryptococcu
727	33	58.9	1483	2	Q5KGL7	CRYNE	Q5kg17	cryptococcu
728	33	58.9	1755	2	Q8CHC7	MOUSE	Q8chc7	mus musculu
729	33	58.9	2000	2	Q61N09	CAEBR	Q61n09	caenorhabdi
730	33	58.9	2068	2	Q57ZG1	9TRYP	Q57zg1	trypanosoma
731	33	58.9	2260	2	Q76KZ5	STRHA	Q76kz5	streptomyce
732	33	58.9	2269	2	Q5SWZ5	MOUSE	Q5swz5	mus musculu
733	33	58.9	2933	2	Q6RKE8	COCHE	Q6rke8	cochliobolu
734	33	58.9	3362	2	Q76KZ4	STRHA	Q76kz4	streptomyce
735	33	58.9	3528	2	Q5Y9G4	9ACTO	Q5y9g4	aeromicrobi
736	33	58.9	5826	2	Q76KY0	STRHA	Q76ky0	streptomyce
737	33	58.9	6521	2	Q8RL72	PSEFL	Q8rl72	pseudomonas
738	32.5	58.0	79	2	Q9AC49	CAUCR	Q9ac49	caulobacter
739	32.5	58.0	88	2	Q914K3	9VIRU	Q914k3	sulfolobus
740	32.5	58.0	130	2	Q8FAB0	ECOL6	Q8fab0	escherichia
741	32.5	58.0	176	2	Q7TVD9	MYCBO	Q7tvd9	mycobacteri
742	32.5	58.0	176	2	Q05443	MYCTU	Q05443	mycobacteri
743	32.5	58.0	188	2	Q4TB10	TETNG	Q4tbi0	tetraodon n
744	32.5	58.0	246	2	Q6LGC3	PHOPR	Q6lgc3	photobacter
745	32.5	58.0	343	2	Q6LRF7	PHOPR	Q6lrf7	photobacter
746	32.5	58.0	345	2	Q6KD97	ECOLI	Q6kd97	escherichia
747	32.5	58.0	345	2	Q8FIW8	ECOL6	Q8fiw8	escherichia
748	32.5	58.0	355	2	Q6LHY5	PHOPR	Q6lhy5	photobacter
749	32.5	58.0	398	2	Q84701	CHLTR	Q84701	chlamydia t
750	32.5	58.0	402	2	Q7VXD6	BORPE	Q7vxd6	bordetella
751	32.5	58.0	402	2	Q7W6N8	BORPA	Q7w6n8	bordetella
752	32.5	58.0	402	2	Q7WHM1	BORBR	Q7whm1	bordetella
753	32.5	58.0	567	1	FIG1	HUMAN	Q96rq9	homo sapien
754	32.5	58.0	575	2	Q4GZN2	HUMAN	Q4gzn2	homo sapien
755	32.5	58.0	575	2	Q4GZN1	HUMAN	Q4gzn1	homo sapien
756	32.5	58.0	588	2	Q8TEM5	HUMAN	Q8tem5	homo sapien
757	32.5	58.0	588	2	Q6P2Q3	HUMAN	Q6p2q3	homo sapien
758	32.5	58.0	919	2	Q6CW36	KLULA	Q6cw36	kluveromyc
759	32.5	58.0	1160	2	Q88D24	PSEPK	Q88dz4	pseudomonas
760	32.5	58.0	1186	2	Q523U8	NOCFA	Q5z3u8	nocardia fa
761	32.5	58.0	1904	2	Q4IL82	GIBZE	Q4il82	gibberella

Q8gpi7	staphylococ	61	2	Q8GPI7_STAAU	57.1	32	762	Q8fcv2	escherichia
Q8nue0	staphylococ	61	2	Q8NUE0_STAAW	57.1	32	763	Q6mqn7	bdellovibri
Q9ac61	staphylococ	61	2	Q9AC61_STAAN	57.1	32	764	Q8x809	escherichia
Q6g5v6	staphylococ	61	2	Q6G5V6_STAAS	57.1	32	765	Q55aa8	dictyosteli
Q6z4t8	oryza sativ	62	2	Q6Z4T8_ORYSA	57.1	32	766	Q7x7q3	oryza sativ
Q5mzj0	synechococc	67	2	Q5MZJ0_SYNP6	57.1	32	767	Q8a2h7	bacteroides
Q9kri1	vibrio chol	73	2	Q9KRI1_VIBCH	57.1	32	768	Q4wi49	aspergillus
Q92k84	rhizobium m	73	2	Q92K84_RHIME	57.1	32	769	Q7S029	neurospora
Q67mp1	symbiobacte	74	2	Q67MP1_SYMTH	57.1	32	770	Q4ZP54	pseudomonas
Pl8646	vigna ungui	75	1	1OKD_VIGUN	57.1	32	771	Q87X45	pseudomonas
Q8cgl7	trichys fas	75	2	Q8CGL7_9HYST	57.1	32	772	Q7UYC9	rhodopirell
Q8ejd1	shewanella	78	2	Q8EJD1_SHEON	57.1	32	773	Q8x701	escherichia
P72271	rhodococcus	85	2	P72271_RHOER	57.1	32	774	Q8cxx5	escherichia
Q81cw7	bacillus ce	87	2	Q81CW7_BACCR	57.1	32	775	P06993	escherichia
Q947f7	glycine max	96	2	Q947F7_SOYBN	57.1	32	776	Q6lhm8	photobacter
Q7tm50	uncultured	96	2	Q7TM50_9ZZZZ	57.1	32	777	Q9ttj4	oryctolagus
Q7tm02	uncultured	103	2	Q7TM02_9ZZZZ	57.1	32	778	Q7zvl4	brachydanio
Q7vk37	helicobacte	105	2	Q7VK37_HELHP	57.1	32	779	Q4szm3	tetraodon n
Q4zyt6	pseudomonas	106	2	Q4ZYT6_PSESY	57.1	32	780	Q4sn87	tetraodon n
Q4iyj6	azotobacter	106	2	Q4IYJ6_AZOVI	57.1	32	781	Q4zp59	pseudomonas
Q5pbt1	anaplasma m	106	2	Q5PBT1_ANAMM	57.1	32	782	Q4nzu4	anaeromyxob
Q73zm0	mycobacteri	110	2	Q73ZM0_MYCPA	57.1	32	783	Q8nit8	neurospora
Q9ala7	synechococc	111	2	Q9ALA7_SYNP7	57.1	32	784	Q6ZR04	homo sapien
Q5n062	synechococc	111	2	Q5N062_SYNP6	57.1	32	785	Q7ZYZ1	brachydanio
Q4kym5	oryza sativ	119	2	Q4KYM5_ORYSA	57.1	32	786	Q4syg9	tetraodon n
Q8wmv6	sus scrofa	124	2	Q8WMV6_PIG	57.1	32	787	Q8c8t6	mus musculu
Q852y8	mycobacteri	124	2	Q852Y8_9CAUD	57.1	32	788	Q86jp2	dictyosteli
Q527j9	magnaporthe	125	2	Q527J9_MAGGR	57.1	32	789	Q4hwp8	gibberella
Q58a12	cucumis mel	131	2	Q58A12_CUCME	57.1	32	790	P40818	homo sapien
Q4kdi0	pseudomonas	134	2	Q4KDI0_PSEF5	57.1	32	791	Q4q685	leishmania
Q7sc72	neurospora	138	2	Q7SC72_NEUCR	57.1	32	792	Q9Y7C1	magnaporthe
Q9h8c4	homo sapien	143	2	Q9H8C4_HUMAN	57.1	32	793	Q51ta9	magnaporthe
Q69k49	oryza sativ	150	2	Q69K49_ORYSA	57.1	32	794	Q76yv0	bacterioph
Q8qrx5	pongine her	156	2	Q8QRX5_9BETA	57.1	32	795	Q5az14	aspergillus
Q9i2i0	pseudomonas	158	2	Q9I2I0_PSEAE	57.1	32	796	Q522r0	magnaporthe
Q6mq88	bdellovibri	160	2	Q6MQ88_BDEBA	57.1	32	797	Q7xh59	oryza sativ
Q5p7t4	azoarcus sp	170	2	Q5P7T4_AZOSE	57.1	32	798	Q8W5K9	oryza sativ
Q4l8m7	staphylococ	171	2	Q4L8M7_STAHJ	57.1	32	799	Q5S864	cryptococcu
Q7uyx1	rhodopirell	171	2	Q7UYX1_RHOBA	57.1	32	800	Q5kg17	cryptococcu
Q88e67	pseudomonas	172	2	Q88E67_PSEPK	57.1	32	801	Q8chc7	mus musculu
Q8kjd2	rhizobium l	174	2	Q8KJD2_RHILO	57.1	32	802	Q61n09	caenorhabdi
Q9zgs2	escherichia	176	2	Q9ZGS2_ECO57	57.1	32	803	Q57zg1	trypanosoma
Q5m9b6	xenopus lae	177	2	Q5M9B6_XENLA	57.1	32	804	Q76kz5	streptomyce
Q6z462	oryza sativ	178	2	Q6Z462_ORYSA	57.1	32	805	Q5swz5	mus musculu
Q4kab0	pseudomonas	179	2	Q4KAB0_PSEF5	57.1	32	806	Q6rke8	cochliobolu
Q6ii35	drosophila	182	2	Q6I135_DROME	57.1	32	807	Q76kz4	streptomyce
Q62373	caenorhabdi	182	2	Q62373_CAEEL	57.1	32	808	Q5y9g4	aeromicrobi
Q9fhl4	arabidopsis	182	2	Q9FHL4_ARATH	57.1	32	809	Q76ky0	streptomyce
Q4miw2	bacillus ce	182	2	Q4MIW2_BACCE	57.1	32	810	Q8rl72	pseudomonas
Q6hi77	bacillus th	182	2	Q6HI77_BACCK	57.1	32	811	Q9ac49	caulobacter
Q737i1	bacillus ce	182	2	Q737I1_BACCI	57.1	32	812	Q914k3	sulfolobus
Q81pz9	bacillus an	182	2	Q81PZ9_BACAN	57.1	32	813	Q8fab0	escherichia
Q6lg71	photobacter	186	2	Q6LG71_PHOPR	57.1	32	814	Q7tvd9	mycobacteri
Q8ys94	anabaena sp	186	2	Q8YS94_ANASP	57.1	32	815	Q05443	mycobacteri
Q64un9	bacteroides	191	2	Q64UN9_BACFR	57.1	32	816	Q4tbi0	tetraodon n
Q9rx50	deinococcus	194	2	Q9RX50_DEIRA	57.1	32	817	Q6lgc3	photobacter
Q5i0s1	xenopus tro	195	2	Q5I0S1_XENTR	57.1	32	818	Q6lrf7	photobacter
Q58353	pyrococcus	197	2	Q58353_PYRHO	57.1	32	819	Q8fiw8	escherichia
Q9p000	homo sapien	198	1	COMD9_HUMAN	57.1	32	820	Q6lhy5	photobacter
Q8k2q0	mus musculu	198	1	COMD9_MOUSE	57.1	32	821	Q84701	chlamydia t
Q53fr9	homo sapien	198	2	Q53FR9_HUMAN	57.1	32	822	Q7vxd6	bordetella
Q5tyh1	anopheles g	200	2	Q5TYH1_ANOGA	57.1	32	823	Q7w6n8	bordetella
Q8gwr6	arabidopsis	203	2	Q8GWR6_ARATH	57.1	32	824	Q7whm1	bordetella
Q6f443	plutella xy	204	2	Q6F443_PLUXY	57.1	32	825	Q7w6n8	bordetella
Q4nwn3	anaeromyxob	205	2	Q4NWN3_9DELT	57.1	32	826	Q96rq9	homo sapien
Q82ps6	streptomyce	205	2	Q82PS6_STRAW	57.1	32	827	Q7whm1	bordetella
Q6zhf1	oryza sativ	206	2	Q6ZHF1_ORYSA	57.1	32	828	Q7w6n8	bordetella
Q7ur11	rhodopirell	206	2	Q7UR11_RHOBA	57.1	32	829	Q7w6n8	bordetella
Q53nc9	oryza sativ	210	2	Q53NC9_ORYSA	57.1	32	830	Q7w6n8	bordetella
Q5n2l3	synechococc	210	2	Q5N2L3_SYNP6	57.1	32	831	Q7w6n8	bordetella
Q677f9	hyacinthus	213	2	Q677F9_9ASPA	57.1	32	832	Q7w6n8	bordetella
Q7z6r3	homo sapien	216	2	Q7Z6R3_HUMAN	57.1	32	833	Q7w6n8	bordetella
Q6a970	propionibac	221	2	Q6A970_PROAC	57.1	32	834	Q7w6n8	bordetella

835 32 57.1 223 2 Q4I2E5_GIBZE Q4i2e5 gibberella
836 32 57.1 223 2 Q5Z5X5_ORYSA Q5z5x5 oryza sativ
837 32 57.1 223 2 Q5YV50_NOCFA Q5yv50 nocardia fa
838 32 57.1 224 2 Q5Zf83_PLAMJ Q5zf83 plantago ma
839 32 57.1 224 2 Q7L1G5_TOBAC Q7l1g5 nicotiana t
840 32 57.1 226 2 Q86SR2_HUMAN Q86sr2 homo sapien
841 32 57.1 228 2 Q9CNM1_PASMU Q9cnm1 pasteurella
842 32 57.1 228 2 Q82LX4_STRAW Q82lx4 streptomyce
843 32 57.1 229 2 Q4J663_AZOVI Q4j663 azotobacter
844 32 57.1 230 2 Q5SSW8_HUMAN Q5ssw8 homo sapien
845 32 57.1 230 2 Q9SIV6_ARATH Q9siv6 arabidopsis
846 32 57.1 230 2 Q9ZB92_RHOER Q9zb92 rhodococcus
847 32 57.1 230 2 Q7NS28_CHRVO Q7ns28 chromobacte
848 32 57.1 237 2 Q5SSX0_HUMAN Q5ssx0 homo sapien
849 32 57.1 239 2 Q6DH54_BRARE Q6dh54 brachydanio
850 32 57.1 243 2 Q5TY12_ANOGA Q5ty12 anopheles g
851 32 57.1 245 2 Q4NI84_9MICC Q4ni84 arthrobacte
852 32 57.1 246 2 Q8Y0L4_RALSO Q8y0l4 ralstonia s
853 32 57.1 247 2 Q4TP77_9SPHN Q4tp77 erythrobact
854 32 57.1 250 2 Q7NG30_GLOVI Q7ng30 gloebacter
855 32 57.1 250 2 Q82PG7_STRAW Q82pg7 streptomyce
856 32 57.1 251 1 CAP7_HUMAN CAP7 homo sapien
857 32 57.1 251 2 Q52LG4_HUMAN Q52lg4 homo sapien
858 32 57.1 253 2 Q6YYI3_ORYSA Q6yyi3 oryza sativ
859 32 57.1 253 2 Q89GZ4_BRAJA Q89gz4 bradyrhizob
860 32 57.1 254 2 Q88YF6_LACPL Q88yf6 lactobacill
861 32 57.1 255 2 Q7SE51_NEUCR Q7se51 neurospora
862 32 57.1 255 2 Q6N9Q3_RHOPA Q6n9q3 rhodopseudo
863 32 57.1 256 2 Q9KX37_9RHIZ Q9kx37 methylocyst
864 32 57.1 256 2 Q7NUF0_CHRVO Q7nuf0 chromobacte
865 32 57.1 260 2 Q568S5_BRARE Q568s5 brachydanio
866 32 57.1 261 2 Q92TL1_RHIME Q9zfe4 pseudomonas
867 32 57.1 265 1 FABI_PSEAE Q5txh7 anopheles g
868 32 57.1 267 2 Q5TXH7_ANOGA Q5txh7 anopheles g
869 32 57.1 268 2 Q52BL3_MAGGR Q52bl3 magnaporthe
870 32 57.1 268 2 P93332_MEDTR P93332 medicago tr
871 32 57.1 270 2 Q4M0V0_9BURK Q4m0v0 burkholderi
872 32 57.1 272 2 Q6X3Z0_CHICK Q6x3z0 gallus gall
873 32 57.1 272 2 Q6X3Z1_CHICK Q6x3z1 gallus gall
874 32 57.1 273 2 Q4V0D2_XANCP Q4v0d2 xanthomonas
875 32 57.1 273 2 Q8PE38_XANCP Q8pe38 xanthomonas
876 32 57.1 278 2 Q6C344_YARLI Q6c344 yarrowia li
877 32 57.1 282 2 Q89HL6_BRAJA Q89hl6 bradyrhizob
878 32 57.1 282 2 Q89PW0_BRAJA Q89pw0 bradyrhizob
879 32 57.1 283 2 Q5M495_STRT2 Q5m495 streptococc
880 32 57.1 284 2 Q56WY4_ARATH Q56wy4 arabidopsis
881 32 57.1 284 2 Q6N287_RHOPA Q6n287 rhodopseudo
882 32 57.1 285 2 Q33629_SACER Q33629 saccharopol
883 32 57.1 285 2 Q4KHK1_PSEF5 Q4khk1 pseudomonas
884 32 57.1 286 2 Q51J97_MAGGR Q51j97 magnaporthe
885 32 57.1 286 2 Q9A6L2_CAUCR Q9a6l2 caulobacter
886 32 57.1 287 2 Q58DH6_BOVIN Q58dh6 bos taurus
887 32 57.1 288 2 Q51SY8_MAGGR Q51sy8 magnaporthe
888 32 57.1 288 2 Q8GHB0_9ACTO Q8ghb0 streptomyce
889 32 57.1 289 2 Q4LT59_9BURK Q4lt59 burkholderi
890 32 57.1 289 2 Q82A27_STRAW Q82a27 streptomyce
891 32 57.1 290 2 Q5CMC7_CRYHO Q5cmc7 cryptospori
892 32 57.1 291 2 Q58AG5_9BURK Q58ag5 ralstonia m
893 32 57.1 293 2 Q7Z5A4_HUMAN Q7z5a4 homo sapien
894 32 57.1 296 2 Q4M0B1_9BURK Q4m0b1 burkholderi
895 32 57.1 296 2 Q7NUS2_CHRVO Q7nus2 chromobacte
896 32 57.1 299 2 Q61CS5_CAEBR Q61cs5 caenorhabdi
897 32 57.1 299 2 Q83B48_COXBU Q83b48 coxiella bu
898 32 57.1 301 2 Q9HS99_HALSA Q9hs99 halobacteri
899 32 57.1 301 2 Q517Z8_ENTHI Q517z8 entamoeba h
900 32 57.1 302 2 Q5H2Y7_XANOR Q5h2y7 xanthomonas
901 32 57.1 308 2 Q7NYJ2_CHRVO Q7nyj2 chromobacte
902 32 57.1 311 2 Q8FWZ7_BRUSU Q8fwz7 brucella su
903 32 57.1 312 2 Q5TRY8_ANOGA Q5try8 anopheles g
904 32 57.1 317 2 Q9RJ64_STRCO Q9rj64 streptomyce
905 32 57.1 318 2 O61826_CAEEL O61826 caenorhabdi
906 32 57.1 320 2 Q5V0P1_HALMA Q5v0p1 haloarcula
907 32 57.1 322 2 O24547_VIGUN O24547 vigna ungui

308 32 57.1 324 2 Q9YEP2_AERPE Q9yep2 aeropyrum p
309 32 57.1 326 2 Q7USG3_RHOBA Q7usg3 rhodopirell
310 32 57.1 326 2 Q5SIG8_THETA Q5sig8 thermus the
311 32 57.1 326 2 Q72IU7_THET2 Q72iu7 thermus the
312 32 57.1 327 2 Q7VZS8_BORPE Q7vzs8 bordetella
313 32 57.1 327 2 Q7W583_BORPA Q7w583 bordetella
314 32 57.1 327 2 Q7WCR3_BORBR Q7wcr3 bordetella
315 32 57.1 327 2 Q60DR3_ORYSA Q60dr3 oryza sativ
316 32 57.1 328 2 Q6A8Q0_PROAC Q6a8q0 propionibac
317 32 57.1 330 2 Q7WBZ8_BORPA Q7wbz8 bordetella
318 32 57.1 330 2 Q7WPZ7_BORBR Q7wpz7 bordetella
319 32 57.1 335 1 SPM1_RAT P56225 rattus norv
320 32 57.1 335 2 Q75QH4_CAPCH Q75qh4 capicum ch
321 32 57.1 337 2 Q8ELP2_OCEIH Q8elp2 oceanobacil
322 32 57.1 340 2 Q4HU90_GIBZE Q4hu90 gibberella
323 32 57.1 340 2 Q8PI52_XANAC Q8pi52 xanthomonas
324 32 57.1 341 2 Q95VR4_TRIVA Q95vr4 trichomonas
325 32 57.1 341 2 Q4HAN3_9DEIO Q4han3 deinococcus
326 32 57.1 344 2 Q4USX3_XANCP Q4usx3 xanthomonas
327 32 57.1 344 2 Q6AL73_DESPS Q6al73 desulfotale
328 32 57.1 344 2 Q7NEH3_GLOVI Q7neh3 gloebacter
329 32 57.1 344 2 Q8PAP2_XANCP Q8pap2 xanthomonas
330 32 57.1 345 2 Q6LIX5_CAEBR Q6lix5 caenorhabdi
331 32 57.1 345 2 Q9CNK9_PASMU Q9cnk9 pasteurella
332 32 57.1 346 2 O24548_VIGUN O24548 vigna ungui
333 32 57.1 347 2 Q7QB73_ANOGA Q7qb73 anopheles g
334 32 57.1 347 2 Q67PH3_SYMTH Q67ph3 symbiobacte
335 32 57.1 349 2 O82705_ORYSA O82705 oryza sativ
336 32 57.1 349 2 Q9RHR2_AZOVI Q9rhr2 azotobacter
337 32 57.1 349 2 Q4IY49_AZOVI Q4iy49 azotobacter
338 32 57.1 350 2 Q5L4X0_CHLAB Q5l4x0 chlamydophi
339 32 57.1 351 2 Q6C6P0_YARLI Q6c6p0 yarrowia li
340 32 57.1 351 2 Q7N5H4_PHOLL Q7n5h4 photorhabdu
341 32 57.1 351 2 Q4RG21_TETNG Q4rg21 tetraodon n
342 32 57.1 352 2 Q8ZKS5_SALTY Q8zks5 salmonella
343 32 57.1 352 2 Q69F96_PHAVU Q69f96 phaseolus v
344 32 57.1 353 2 Q4KKQ8_PSEF5 Q4kkq8 pseudomonas
345 32 57.1 353 2 Q8MHQ3_CHICK Q8mhq3 gallus gall
346 32 57.1 353 2 Q9BCW3_GALLS Q9bcw3 gallus gall
347 32 57.1 355 2 Q5SKQ8_THET8 Q5skq8 thermus the
348 32 57.1 355 2 Q72L47_THET2 Q72l47 thermus the
349 32 57.1 357 2 O22427_ARATH O22427 arabidopsis
350 32 57.1 357 2 Q9SRT9_ARATH Q9srt9 arabidopsis
351 32 57.1 357 2 Q50HV8_WHEAT Q50hv8 triticum ae
352 32 57.1 357 2 Q7XB44_HORVU Q7xb44 hordeum vul
353 32 57.1 357 2 Q4KE51_PSEF5 Q4ke51 pseudomonas
354 32 57.1 359 2 Q8VWN8_GOSHI Q8vwn8 gossypium h
355 32 57.1 360 2 Q8LB19_ARATH Q8lb19 arabidopsis
356 32 57.1 360 2 Q9LFW1_ARATH Q9lfw1 arabidopsis
357 32 57.1 360 2 Q4LZF0_9BURK Q4lzf0 burkholderi
358 32 57.1 360 2 Q82H28_STRAW Q82h28 streptomyce
359 32 57.1 360 2 Q89ID2_BRAJA Q89id2 bradyrhizob
360 32 57.1 361 2 Q762H2_RHORH Q762h2 rhodococcus
361 32 57.1 361 2 Q4NJZ7_9MICC Q4njz7 arthrobacte
362 32 57.1 362 2 O22666_ARATH O22666 arabidopsis
363 32 57.1 362 2 Q6IV07_LYCES Q6iv07 lycopersico
364 32 57.1 362 2 Q9XE81_SORBI Q9xe81 sorghum bic
365 32 57.1 363 2 Q6MBF3_PARUW Q6mbf3 parachlamyd
366 32 57.1 364 1 UPTG_MAIZE P80607 zea mays (m
367 32 57.1 364 1 UPTG_PEA O04300 pisum sativ
368 32 57.1 364 2 O22428_ARATH O22428 arabidopsis
369 32 57.1 364 2 Q8H8T0_ORYSA Q8h8t0 oryza sativ
370 32 57.1 364 2 Q9FUN9_ORYSA Q9fun9 oryza sativ
371 32 57.1 364 2 Q9ZR35_ORYSA Q9zr35 oryza sativ
372 32 57.1 364 2 Q8D549_VIBVU Q8d549 vibrio vuln
373 32 57.1 364 2 Q7MBY6_VIBVY Q7mby6 vibrio vuln
374 32 57.1 365 1 UPTG1_SOLTU Q9sc19 solanum tub
375 32 57.1 365 2 Q51LV9_MAGGR Q51lv9 magnaporthe
376 32 57.1 366 1 UPTG2_SOLTU Q8ru27 solanum tub
377 32 57.1 366 2 Q6Z4G3_ORYSA Q6z4g3 oryza sativ
378 32 57.1 366 2 Q57HH7_SALCH Q57hh7 salmonella
379 32 57.1 366 2 O84715_CHLTR O84715 chlamydia t
380 32 57.1 366 2 Q5PKF8_SALPA Q5pkf8 salmonella

981 32 57.1 366 2 Q821M3 CHLCV Q821m3 chlamydia p
982 32 57.1 366 2 Q9PLL7_CHLMU Q9pl17 chlamydia m
983 32 57.1 366 2 Q8Z2U9_SALTI Q8z2u9 salmonella
984 32 57.1 366 2 Q89Y61_BRAJA Q89y61 bradyrhizob
985 32 57.1 366 2 Q9Z756_CHLPN Q9z756 chlamydia p
986 32 57.1 368 2 Q9ZR33_WHEAT Q9zr33 triticum ae
987 32 57.1 368 2 Q5H1V8_XANOR Q5hlv8 xanthomonas
988 32 57.1 368 2 Q8R1X0_MOUSE Q8rlx0 mus musculus
989 32 57.1 369 2 Q5TTN8_ANOGE Q5ttn8 anopheles g
990 32 57.1 370 2 Q9I410_PSEAE Q9i410 pseudomonas
991 32 57.1 370 2 Q9RRJ7_DEIRA Q9rrj7 deinococcus
992 32 57.1 370 2 Q83605_TREPA Q83605 treponema p
993 32 57.1 371 2 Q81753_ARATH Q81753 arabidopsis
994 32 57.1 372 2 Q5TS89_ANOGA Q5ts89 anopheles g
995 32 57.1 372 2 Q60B43_METCA Q60b43 methylococc
996 32 57.1 373 2 Q9SR90_ARATH Q9sr90 arabidopsis
997 32 57.1 376 2 Q5ENA4_9BACT Q5ena4 uncultured
998 32 57.1 377 2 Q55W41_CRYNE Q55w41 cryptococcu
999 32 57.1 377 2 Q5KK53_CRYNE Q5kk53 cryptococcu
1000 32 57.1 378 2 Q9SUE4_ARATH Q9sue4 arabidopsis

ALIGNMENTS

RESULT 1
Q4JY45_CORJK
ID Q4JY45_CORJK PRELIMINARY; PRT; 242 AA.
AC Q4JY45;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE MsrA protein (EC 1.8.4.6).
GN Name=msrA; ORFNames=jk0110;
OS Corynebacterium jeikeium (strain K411).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=306537;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
RA Albersmeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
RA Kalinowski J., Meyer F., Rupp O., Schneiker S., Viehoveer P.,
RA Puhler A.;
RT "Complete Genome Sequence and Analysis of the Multiresistant
RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requiring
RT Bacterium of the Human Skin Flora.";
RL J. Bacteriol. 187:4671-4682(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RA Linke B., Tauch A.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR931997; CAI36262.1; -; Genomic_DNA.
KW Oxidoreductase.
SQ SEQUENCE 242 AA; 26352 MW; B1776A45CDAE8D3E CRC64;

Query Match 76.8%; Score 43; DB 2; Length 242;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||||:||||
Db 171 GRVQAIAAGW 180

RESULT 2
Q8RSP4_RHILV
ID Q8RSP4_RHILV PRELIMINARY; PRT; 467 AA.
AC Q8RSP4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FixW3 protein.
GN Name=fixW3;
OS Rhizobium leguminosarum (biovar viciae).
OG Plasmid pRL6JI.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15375140; DOI=10.1128/JB.186.19.6586-6594.2004;
RA Martinez M., Palacios J.M., Imperial J., Ruiz-Argueso T.;
RT "Symbiotic autoregulation of nifA expression in Rhizobium
RT leguminosarum bv. viciae.";
RL J. Bacteriol. 186:6586-6594(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Martinez M.;
RT "Regulation simbiotica de la expresion del sistema hidrogenasa por
RT NifA en R. leguminosarum bv. viciae.";
RL Thesis (2000), Department of Biotechnology, Universidad Politecnica
RL Madrid, Madrid, Spain.
DR EMBL; AJ431176; CAD24022.1; -; Genomic_DNA.
KW Plasmid.
SQ SEQUENCE 467 AA; 52042 MW; 412603D21D20AAA7 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 467;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||||:||||
Db 323 GRMRDMDAGW 332

RESULT 3
Q4TM05_9SPHN
ID Q4TM05_9SPHN PRELIMINARY; PRT; 776 AA.
AC Q4TM05;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative cation transporting P-type ATPase.
GN ORFNames=ELI2122;
OS Erythrobacter litoralis HTCC2594.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Erythrobacteraceae; Erythrobacter.
OX NCBI_TaxID=314225;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HTCC2594;
RA Giovannoni S.J., Cho J.-C., Ferreira S., Johnson J., Kravitz S.,
RA Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
RA Friedman R., Venter J.C.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; AAGG01000006; EAL74862.1; -; Genomic_DNA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB_hvy.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR001756; Cu_ATPase.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR007029; YHS.
DR InterPro; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR Pfam; PF04945; YHS; 1.

DR PRINTS; PRO0119; CATATPASE.
DR PRINTS; PRO0943; CUATPASE.
DR TIGRFAMS; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMS; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMS; TIGR01494; ATPase P-type; 2.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN 1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 776 AA; 81530 MW; 22B466303B32A446 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 776;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: |||||
Db 371 RLADQVAGW 379

RESULT 4
P74512 SYN3
ID P74512_SYN3 PRELIMINARY; PRT; 780 AA.
AC P74512;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Cation-transporting ATPase; E1-E2 ATPase.
GN OrderedLocusNames=slr1950;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Saito M., Saito M., Saito M.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; BA000022; BAAL8616.1; -; Genomic_DNA.
DR PIR; S76487; S76487.
DR HSSP; O32220; 10Q3.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004008; F:copper-exporting ATPase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0006825; P:copper ion transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB_hvy.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR001877; Cu ATPase1.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR006121; HeavyMe transpt.
DR InterPro; IPR000150; Hypothet_cof.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PRO0119; CATATPASE.
DR PRINTS; PRO0942; CUATPASEI.
DR TIGRFAMS; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMS; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMS; TIGR01494; ATPase P-type; 2.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN 1.

DR PROSITE; PS01229; COF 2; UNKNOWN 1.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS00846; HMA_2; 1.
KW Complete proteome.
SQ SEQUENCE 780 AA; 82571 MW; 5BC2B2573124708C CRC64;

Query Match 73.2%; Score 41; DB 2; Length 780;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: |||||
Db 357 RLADQVAGW 365

RESULT 5
Q89WI7 BRAJA
ID Q89WI7_BRAJA PRELIMINARY; PRT; 823 AA.
AC Q89WI7;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocusNames=bl10700;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; BA000040; BAC45965.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004008; F:copper-exporting ATPase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0006825; P:copper ion transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB_hvy.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR001756; Cu ATPase.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR012348; Ribncl_rdnase_rel.
DR InterPro; IPR011017; TRASH.
DR InterPro; IPR007029; YHS.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR Pfam; PF04945; YHS; 1.
DR PRINTS; PRO0119; CATATPASE.
DR PRINTS; PRO0943; CUATPASE.
DR SMART; SM00746; TRASH; 1.
DR TIGRFAMS; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMS; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMS; TIGR01494; ATPase P-type; 1.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN 1.

DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE No apparent S. cerevisiae ortholog, putative.
GN Name=PY02659;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteu M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.P., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000732; EAA22136.1; -; Genomic_DNA.
DR InterPro; IPR003639; Mov34-1.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34-1; 1.
SQ SEQUENCE 302 AA; 34663 MW; 3A3B7B0A0E0C2AA1 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 302;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 97 RPRDQVVGW 105
| | | | | | | |

RESULT 10
Q4Z1J9_PLABE PRELIMINARY; PRT; 313 AA.
ID Q4Z1J9_PLABE PRELIMINARY;
AC Q4Z1J9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PB000359.01.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAA101001124; CAH95839.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 36057 MW; 8D08B882139C8CB8 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 313;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 107 RPRDQVVGW 115
| | | | | | | |

RESULT 11
Q4XWR8_PLACH PRELIMINARY; PRT; 315 AA.
ID Q4XWR8_PLACH PRELIMINARY;
AC Q4XWR8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PC001205.02.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAJ01002585; CAH78643.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 36144 MW; B12CA075E710D75A CRC64;

Query Match 71.4%; Score 40; DB 2; Length 315;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 107 RPRDQVVGW 115
| | | | | | | |

RESULT 12
Q8I2X0_PLAF7 PRELIMINARY; PRT; 320 AA.
ID Q8I2X0_PLAF7 PRELIMINARY;
AC Q8I2X0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PF10895c.
GN Name=PF10895c;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Corton C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Doggett J.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabbinoiwitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL929357; CAD51865.1; -; Genomic_DNA.
DR InterPro; IPR003639; Mov34-1.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR ProDom; PD363422; Mov34-1; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Hypothetical protein.
SQ SEQUENCE 320 AA; 36820 MW; 158AA64F9BC5F8A4 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 320;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db 107 RPRDQVVGW 115

RESULT 13
Q7RXX6_NEUCR PRELIMINARY; PRT; 369 AA.
AC Q7RXX6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU00469.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels P., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseelis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seller S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000735; EAA27552.1; -; Genomic_DNA.
DR InterPro; IPR000408; Reg_chr_condens.
DR Pfam; PF00415; RCC1; 1.
DR PRINTS; PR00633; RCCNDNSATION.
DR PROSITE; PS50012; RCC1_3; 4.
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 38697 MW; 0D83BACF5E420C20 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 369;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 245 GEVPDQVVGW 254

RESULT 14
Q7TLZ4_9ZZZZ PRELIMINARY; PRT; 98 AA.
ID Q7TLZ4_9ZZZZ
AC Q7TLZ4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative xylanase (Fragment).
OS uncultured organism.
OC unclassified; environmental samples.
OX NCBI_TaxID=155900;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Radomski C.C.A., Seow K.T., Warren R.A.J., Yap W.H.;
RT "Method for isolating xylanase gene sequences from soil DNA,
RT compositions useful in such method and compositions obtained
RT thereby.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Chitteranjan S., Radomski C.C., Chow M.L., Davies J., Axelrood P.E.;
RT "Retrieval of novel DNA sequences encoding xylanase-like genes from
RT British Columbia forest soils.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427439; AAP87537.1; -; Genomic_DNA.
DR HSP; P26514; IEOW.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001000; Glyco_hydro_10.
DR Pfam; PF00331; Glyco_hydro_10; 1.
KW Glycosidase; Hydrolase; Xylan degradation.
FT NON_TER 1 1
FT NON_TER 98 98
SQ SEQUENCE 98 AA; 10854 MW; 4142C83DD2E89A70 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 98;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 38 GRYGRGIAGW 47

RESULT 15
Q9FRGL_ORYSA PRELIMINARY; PRT; 124 AA.
ID Q9FRGL_ORYSA PRELIMINARY;
AC Q9FRGL;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Expressed protein.
GN Name=Os03g47210;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBb0009F04 genomic sequence.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

RA Buell R.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079830; AAG46067.1; -; Genomic_DNA.
DR Gramene; Q9FRG1; -.
SQ SEQUENCE 124 AA; 12612 MW; B7AACB98AC6C2E78 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 124;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| |:
Db 35 GRPLDRVAGW 44

RESULT 16
HIS1 CAUCR
ID HIS1 CAUCR STANDARD; PRT; 320 AA.
AC Q9A2P5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ATP phosphoribosyltransferase (EC 2.4.2.17) (ATP-PRTase) (ATP-PRT).
GN Name=hisG; OrderedLocusNames=CC3511;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- FUNCTION: Catalyzes the condensation of ATP and PRPP to form N'-
5'-phosphoribosyl-ATP (PR-ATP). Has a crucial role in the pathway
CC because the rate of histidine biosynthesis seems to be controlled
CC primarily by regulation of hisG enzymatic activity (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribosyl)-ATP + diphosphate =
CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Feedback inhibited by histidine (By
CC similarity).
CC -!- PATHWAY: Amino-acid biosynthesis; L-histidine biosynthesis; L-
CC histidine from PRPP: step 1.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.
CC Long subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

EMBL; AE006009; AAK25473.1; -; Genomic_DNA.
DR PIR; E87684; E87684.
DR TIGR; CC3511; -.
DR HAMAP; MF 00079; -; 1.
DR InterPro; IPR001348; ATP_phospho_trans.
DR Pfam; PF01634; HisG; 1.
DR ProDom; PD003516; ATP_phospho_trans; 1.
DR PROSITE; PS01316; ATP_P_PHORIBOSYLTR; 1.
KW Amino-acid biosynthesis; Complete proteome; Glycosyltransferase;

KW Histidine biosynthesis; Magnesium; Metal-binding; Transferase.
SQ SEQUENCE 320 AA; 33587 MW; B3126BC54400A6C1 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 320;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| |:
Db 13 GRLKDQVEAW 22

RESULT 17
Q4LRX7 9BURK
ID Q4LRX7_9BURK PRELIMINARY; PRT; 334 AA.
AC Q4LRX7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Regulatory protein, LysR:LysR, substrate-binding.
GN ORFNames=Bcen2424DRAFT_3611;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHL01000022; EAM18796.1; -; Genomic_DNA.
SQ SEQUENCE 334 AA; 35491 MW; OD08091E13868B4E CRC64;

Query Match 69.6%; Score 39; DB 2; Length 334;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||| |:
Db 254 GRLREVLAGW 263

RESULT 18
Q8YY68 ANASP
ID Q8YY68 ANASP PRELIMINARY; PRT; 360 AA.
AC Q8YY68;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE All0983 protein.
GN OrderedLocusNames=all0983;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

RA	Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,	RA	Blackstock W.P., Choudhary J.S., Grant S.G.;
RA	Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,	RT	"Proteomic analysis of in vivo phosphorylated synaptic proteins.";
RA	Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,	RL	J. Biol. Chem. 280:5972-5982(2005).
RA	Yasuda M., Tabata S.;	CC	- - FUNCTION: Delays the onset of cell spreading in the early stages
RT	"Complete genomic sequence of the filamentous nitrogen-fixing	CC	of cell adhesion to fibronectin. Also involved in calcium-
RT	Cyanobacterium Anabaena sp. strain PCC 7120.";	CC	dependent exocytosis from PC12 cells (By similarity).
RL	DNA Res. 8:205-213(2001).	CC	- - SUBUNIT: Interacts with BCAR1/p130Cas through its C-terminal
DR	EMBL; BA000019; BAB72940.1; -; Genomic_DNA.	CC	domain. Also interacts with the N-terminal coiled-coil region of
DR	PIR; AD1929; AD1929.	CC	SNAP25 (By similarity).
DR	HSSP; P05806; INPC.	CC	- - ALTERNATIVE PRODUCTS:
DR	GO; GO:0005576; C:extracellular region; IEA.	CC	Event=Alternative splicing; Named isoforms=3;
DR	GO; GO:0004222; F:metalloendopeptidase activity; IEA.	CC	Name=1; Synonyms=1a;
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	CC	IsolId=Q9QW16-1; Sequence=Displayed;
DR	InterPro; IPR001570; Peptidase_M4.	CC	Name=2; Synonyms=1b;
DR	InterPro; IPR006025; Pept M Zn_BS.	CC	IsolId=Q9QW16-2; Sequence=VSP_050631;
DR	Pfam; PF01447; Peptidase_M4; 1.	CC	Name=3;
DR	Pfam; PF02868; Peptidase_M4 C; 1.	CC	IsolId=Q9QW16-3; Sequence=VSP_050632;
DR	PRINTS; PR00730; THERMOLYSIN.	CC	Note=No experimental confirmation available;
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.	CC	- - TISSUE SPECIFICITY: Expressed mainly in brain, testis and
KW	Complete proteome.	CC	epithelial-rich tissues such as mammary gland, lung and kidney.
SQ	SEQUENCE 360 AA; 39795 MW; D14F84AA488CA6F1 CRC64;	CC	- - PTM: Tyrosine-phosphorylated in response to EGF and to cell
		CC	adhesion to integrin ligands (By similarity).
		CC	- - SIMILARITY: Belongs to the p140Cap family.
		CC	- - CAUTION: Ref.2 sequence differs from that shown due to frameshifts
		CC	in positions 483 and 527.
		CC	-----
QY	1 GRVRDQVAGW 10	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
Db	265 GHVKDQYTGW 274	CC	the European Bioinformatics Institute. There are no restrictions on its
		CC	use as long as its content is in no way modified and this statement is not
		CC	removed.
		CC	-----
RESULT 19		CC	EMBL; AF040944; AAC15635.1; -; mRNA.
SNIP_MOUSE		DR	EMBL; U59873; AAD00087.1; -; mRNA.
ID_SNIP_MOUSE	STANDARD; PRT; 1250 AA.	DR	EMBL; AK129422; BAC98232.1; ALT_FRAME; mRNA.
AC	Q9QW16; O70298;	DR	PIR; T34101; T34101.
DT	29-MAR-2004 (Rel. 43, Created)	DR	Ensembl; ENSMUSG00000038453; Mus musculus.
DT	29-MAR-2004 (Rel. 43, Last sequence update)	DR	MGI; MGI:1933179; P140.
DT	13-SEP-2005 (Rel. 48, Last annotation update)	DR	GO; GO:0005515; F:protein binding; IPI.
DE	p130Cas-associated protein (p140Cap) (SNAP-25-interacting protein)	DR	GO; GO:0007162; P:negative regulation of cell adhesion; IDA.
DE	(SNIP).	DR	GO; GO:0007162; P:negative regulation of cell adhesion; IDA.
GN	Name=P140; Synonyms=Kiaa1684;	KW	Alternative splicing; Coiled coil; Cytoskeleton; Exocytosis;
OS	Mus musculus (Mouse).	KW	Phosphorylation.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	FT	REGION 681 731 Interaction with SNAP25 (By similarity).
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;	FT	COILED 630 654 Potential.
OC	Muroidea; Muridae; Murinae; Mus.	FT	COILED 713 845 Potential.
OX	NCBI_TaxID=10090;	FT	COILED 994 1011 Potential.
RN	[1]	FT	MOD RES 1054 1054 Phosphoserine.
RP	NUCLEOTIDE SEQUENCE (ISOFORMS 1 AND 2).	FT	VARSP LIC 1 41 MQPWQCLRRFALAWWERTAEGRARSPREEVGPDRDPGGRGEP
RA	Croci L., Bossolasco M., Consalez G.G.;	FT	Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RL	[2]	FT	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 271-1250 (ISOFORM 3).
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 271-1250 (ISOFORM 3).	RC	TISSUE=Brain;
RC	TISSUE=Brain;	RX	PubMed=14621295;
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,	RA	Saga Y., Nagase T., Ohara O., Koga H.;
RT	"Prediction of the coding sequences of mouse homologues of KIAA gene:	RT	III. The complete nucleotide sequences of 500 mouse KIAA-homologous
RT	cDNAs identified by screening of terminal sequences of cDNA clones	RT	randomly sampled from size-fractionated libraries.";
RL	DNA Res. 10:167-180(2003).	RL	[3]
RN	[3]	RP	ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RP	PubMed=14657239; DOI=10.1091/mbc.E03-09-0689;	RX	Di Stefano P., Cabodi S., Erba E.B., Margaria V., Bergatto E.,
RA	Giuffrida M.G., Silengo L., Tarone G., Turco E., Defilippi P.;	RA	"p130Cas-associated protein (p140Cap) as a new tyrosine-phosphorylated
RT	protein involved in cell spreading.";	RT	Mol. Biol. Cell 15:787-800(2004).
RL	[4]	RN	[4]
RP	PHOSPHORYLATION SITE SER-1054.	RP	PHOSPHORYLATION SITE SER-1054.
RX	PubMed=15572359; DOI=10.1074/jbc.M411220200;	RX	Collins M.O., Yu L., Caba M.P., Husi H., Campuzano I.,
RA		RA	

Query Match 69.6%; Score 39; DB 1; Length 1250;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|
Db 138 RMREQVGW 146

RESULT 20
Q4QDJ2 LEIMA
ID Q4QDJ2_LEIMA PRELIMINARY; PRT; 1370 AA.
AC Q4QDJ2;

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohamoud Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU48462.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 233 AA; 24465 MW; FF961AE6D98CF48C CRC64;

Query Match 67.9%; Score 38; DB 2; Length 233;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db 201 RLRAQIAGW 209
|:|:|:|:|
|:|:|:|:|

RESULT 24
Q63VC8 BURPS
ID Q63VC8 BURPS PRELIMINARY; PRT; 234 AA.
AC Q63VC8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BPSL1314;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songsiwilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH35311.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 234 AA; 24596 MW; 8237542235E1815B CRC64;

Query Match 67.9%; Score 38; DB 2; Length 234;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db 202 RLRAQIAGW 210
|:|:|:|:|
|:|:|:|:|

RESULT 25
Q4WRF3 ASPFU
ID Q4WRF3 ASPFU PRELIMINARY; PRT; 243 AA.
AC Q4WRF3;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Heme/steroid binding domain protein, putative.
GN ORFNames=Afulg16510;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majeros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penalva M.A., Perteau M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF01000004; EAL90979.1; -; Genomic_DNA.
SQ SEQUENCE 243 AA; 27517 MW; 0DF9E6357B6B37D1 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 243;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 189 GKRKQVAHW 198
|:|:|:|:|
|:|:|:|:|

RESULT 26
DHAA MYCBO
ID DHAA MYCBO STANDARD; PRT; 300 AA.
AC Q9XB14; Q6EUU9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Haloalkane dehalogenase (EC 3.8.1.5).
GN Name=dhaA; Synonyms=dmbA; OrderedLocusNames=Mb2610;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5033/66, and MUI1;
RA Jesenska A., Strouhal M., Pavlova M., Tesinska I., Monincova M.,
RA Bartos M., Pavlik I., Rychlik I., Nagata Y., Damborsky J.;
RT "Mycobacterial haloalkane dehalogenases: cloning, biochemical
RT properties and distribution";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- FUNCTION: Catalyzes hydrolytic cleavage of carbon-halogen bonds in
CC halogenated aliphatic compounds, leading to the formation of the
CC corresponding primary alcohols, halide ions and protons (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 1-haloalkane + H(2)O = a primary alcohol +
CC halide.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: Belongs to the haloalkane dehalogenase family. Type 2
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ243259; CAB45532.2; -; Genomic_DNA.
CC EMBL; AJ784272; CAH04659.1; -; Genomic_DNA.
CC EMBL; BX248343; CAD94795.1; -; Genomic_DNA.
CC HSSP; P51698; 1CV2.
CC HAMAP; MF_01231; -; 1.
CC InterPro; IPR000073; A/b_hydrolase.
CC InterPro; IPR000639; Epox_hydrolase.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00561; Abhydrolase_1; 1.
CC PRINTS; PR00412; EPOXHYDRLASE.
CC Complete proteome; Hydrolase.
KW ACT_SITE 109 Nucleophile (By similarity).
FT ACT_SITE 133 Proton donor (By similarity).
FT ACT_SITE 273 Proton acceptor (By similarity).
FT ACT_SITE 96 T -> A (in Ref. 1).
FT CONFLICT 96
SQ SEQUENCE 300 AA; 33730 MW; BCPA71EC43529F7B CRC64;

Query Match 67.9%; Score 38; DB 1; Length 300;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 252 GRIRDYVRSW 261
||:|||||

RESULT 27
DHAA MYCTU
ID DHAA MYCTU STANDARD; PRT; 300 AA.
AC Q50642;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Haloalkane dehalogenase 3 (EC 3.8.1.5).
GN Name=dhaA; OrderedLocusNames=Rv2579, MT2656; ORFNames=MTCY227.22c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/311159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaiia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,

RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- FUNCTION: Catalyzes hydrolytic cleavage of carbon-halogen bonds in
CC halogenated aliphatic compounds, leading to the formation of the
CC corresponding primary alcohols, halide ions and protons (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 1-haloalkane + H(2)O = a primary alcohol +
CC halide.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: Belongs to the haloalkane dehalogenase family. Type 2
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX842580; CAES5500.1; -; Genomic_DNA.
CC EMBL; AE000516; AAK46969.1; -; Genomic_DNA.
CC PIR; B70725; B70725.
CC HSSP; P51698; 1IZ7.
CC TIGR; MT2656; -.
CC TubercuList; Rv2579; -.
CC HAMAP; MF_01231; -; 1.
CC InterPro; IPR000073; A/b_hydrolase.
CC InterPro; IPR000639; Epox_hydrolase.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00561; Abhydrolase_1; 1.
CC PRINTS; PR00412; EPOXHYDRLASE.
CC Complete proteome; Hydrolase.
KW ACT_SITE 109 109 Nucleophile (By similarity).
FT ACT_SITE 133 133 Proton donor (By similarity).
FT ACT_SITE 273 273 Proton acceptor (By similarity).
SQ SEQUENCE 300 AA; 33728 MW; E8F4854749F22562 CRC64;

Query Match 67.9%; Score 38; DB 1; Length 300;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 252 GRIRDYVRSW 261
||:|||||

RESULT 28
Q7UV53 RHOB
ID Q7UV53 RHOB PRELIMINARY; PRT; 349 AA.
AC Q7UV53;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB2871;
OS Rhodopirellula baltica.

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlessner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294137; CAD72873.1; -; Genomic_DNA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR010372; DNA_pol3_delta.
DR InterPro; IPR005790; DNA_polIII_delta.
DR Pfam; PF06144; DNA_pol3_delta; 1.
DR TIGRFAMs; TIGR01128; hoIA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 349 AA; 38362 MW; 5AC4590EB5845D82 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 349;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
DB 206 VRDVVAGW 213

RESULT 29
Q4H705_9DEIO
ID Q4H705_9DEIO PRELIMINARY; PRT; 367 AA.
AC Q4H705;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase M20:Peptidase M20.
GN ORFNames=DgeodRAFT_0250;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Iarani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000011; EAL82057.1; -; Genomic DNA.
SQ SEQUENCE 367 AA; 39038 MW; 84B167B157FC3CF8 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 367;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10

Db 32 GRVMDVVEGW 41
||| | | | |
RESULT 30
Q89SF5_BRAJA
ID Q89SF5_BRAJA PRELIMINARY; PRT; 367 AA.
AC Q89SF5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN OrderedLocusNames=bl12445;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC47710.1; -; Genomic_DNA.
DR HSSP; Q51990; 1GWJ.
DR GO; GO:0016491; F:Oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001155; Oxidored_FMN.
DR Pfam; PF00724; Oxidored_FMN; 1.
KW Complete proteome.
SQ SEQUENCE 367 AA; 39432 MW; 791C46C4EE39BE4C CRC64;

Query Match 67.9%; Score 38; DB 2; Length 367;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
DB 79 KDQVAGW 85
: | | | | |
||| | | | |
RESULT 31
Q747F0_GEOSL
ID Q747F0_GEOSL PRELIMINARY; PRT; 369 AA.
AC Q747F0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=GSU3317;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Davidsen T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969(2003).

```
DR EMBL; AE017180; AAR36707.1; -; Genomic_DNA.
DR TIGR; GSU3317; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 369 AA; 37161 MW; 1A22761D0A1DAA7D CRC64;

Query Match 67.9%; Score 38; DB 2; Length 369;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|::|:|
Db 80 GKIRGPIAGW 89

RESULT 32
Q9L144_STRCO
ID Q9L144_STRCO PRELIMINARY; PRT; 385 AA.
AC Q9L144;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative racenase.
GN OrderedLocusNames=SCO7415; ORFNames=SC6D11.11;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939131; CAB76334.1; -; Genomic_DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001354; MR_MLE.
DR Pfam; PF01188; MR_MLE; 1.
DR Pfam; PF02746; MR_MLE_N; 1.
KW Complete proteome.
SQ SEQUENCE 385 AA; 41747 MW; 9777B77BA42EB0B7 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 385;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|
Db 159 RLTDQLAGW 167

RESULT 33
Q74SA4_YERPE
ID Q74SA4_YERPE PRELIMINARY; PRT; 443 AA.
AC Q74SA4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Signal transduction histidine kinase.
GN Name=baeS8; OrderedLocusNames=YP2718;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
```

```
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
avirulent to humans.";
RL DNA Res. 11:179-197(2004).
DR EMBL; AE017137; AAS62907.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008358; Lantibiot_regn.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PRINTS; PR01780; LANTIREGPROT.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase.
SQ SEQUENCE 443 AA; 49971 MW; 026FA684A5751F4B CRC64;

Query Match 67.9%; Score 38; DB 2; Length 443;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|
Db 113 KVNDQVVGW 121

RESULT 34
Q66HY4_BRARE
ID Q66HY4_BRARE PRELIMINARY; PRT; 448 AA.
AC Q66HY4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:92214.
GN ORFNames=zgc:92214;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinaki M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; BC081623; AAH81623.1; -; mRNA.
DR Ensembl; ENSDARG0000036491; Danio rerio.
DR ZFIN; ZDB-GENE-040912-153; zgc:92214.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00465; EP450IV.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 448 AA; 51781 MW; 342AB9E9856CBDB5 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 448;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|||:|:|
Db 193 RVKDNISGW 201

RESULT 35
Q668C3 YERPS
ID Q668C3_YERPS PRELIMINARY; PRT; 461 AA.
AC Q668C3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Sensor protein baes (EC 2.7.3.-).
GN Name=baes; OrderedLocusNames=YPTB2817;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH22055.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008358; Lantibiot_regn.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PRINTS; PR01780; LANTIREGPROT.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 461 AA; 51937 MW; 941A4EEE177AD4F4 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 461;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
:|:|:|:|
Db 131 KVNDQVVGW 139

RESULT 36
Q7X541_9ACTO
ID Q7X541_9ACTO PRELIMINARY; PRT; 588 AA.
AC Q7X541;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Acyl-CoA synthase.
GN Name=lipA;
OS Actinoplanes friuliensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micromonosporineae; Micromonosporaceae; Actinoplanes.
OX NCBI_TaxID=196914;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heinzelmann E., Berger S., Muller C., Hartner T., Poralla K.,
RA Wohlleben W., Schwartz D.;
RT "An acyl-CoA dehydrogenase is involved in the formation of the cis3
RT doublebond in the acyl residue of the lipopeptide antibiotic
RT flubimicin in Actinoplanes friuliensis.";
RL Microbiol. 151:1963-1974(2005).
DR EMBL; AJ488769; CAD32910.2; -; Genomic_DNA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
SQ SEQUENCE 588 AA; 62629 MW; AB80EEC26969D107 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 588;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|:|:|
Db 124 GESRDEVTGW 133

RESULT 37
Q869Y0_DICDI
ID Q869Y0_DICDI PRELIMINARY; PRT; 941 AA.
AC Q869Y0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to plasmodium falciparum. cell differentiation protein rccl,

DE putative.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AX4;
RC MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RX Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116305; AAO52297.1; -; Genomic_DNA.
DR InterPro; IPR007216; Rcd1.
DR PANTHER; PTHR12262; Rcd1; 1.
DR Pfam; PF04078; Rcd1; 1.
SQ SEQUENCE 941 AA; 107408 MW; E01F97A1CFF17CA7 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 941;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db :|||: ||
740 KIRDQLIGW 748

RESULT 38
Q553Y8 D1CDI
ID Q553Y8_D1CDI PRELIMINARY; PRT; 941 AA.
AC Q553Y8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cell differentiation family, Rcd1-like protein.
GN Name=DDB0220000; ORFNames=DDB0220000;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulseghe H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF101000028; EAL69925.1; -; Genomic_DNA.

DR InterPro; IPR007216; Rcd1.
DR Pfam; PF04078; Rcd1; 1.
KW Nucleotide-binding.
SQ SEQUENCE 941 AA; 107382 MW; 5D36FB71CFE71079 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 941;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
Db :|||: ||
740 KIRDQLIGW 748

RESULT 39
Q8UE69 AGRT5
ID Q8UE69_AGR5 PRELIMINARY; PRT; 63 AA.
AC Q8UE69;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Atul892.
GN OrderedLocusNames=Atul892;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009143; AAL42888.1; -; Genomic_DNA.
DR PIR; AB2809; AB2809.
DR InterPro; IPR010710; DUF1289.
DR Pfam; PF06945; DUF1289; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 63 AA; 7083 MW; C614BB7BEDC5F4E4 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 63;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db :|||: ||
22 GRIUDEIGGW 31

RESULT 40
Q4LUS0_9BURK
ID Q4LUS0_9BURK PRELIMINARY; PRT; 121 AA.
AC Q4LUS0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Response regulator receiver.
GN ORFNames=Bcen2424DRAFT_4960;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;

```
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Picluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHL01000013; EAM19906.1; -; Genomic DNA.
SQ SEQUENCE 121 AA; 12687 MW; 5E592333A7FB98D4 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 121;
Best Local Similarity 70.0%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
DB 11 GTVRDEVAGF 20

RESULT 41
Q606W1_METCA PRELIMINARY; PRT; 178 AA.
AC Q606W1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Conserved domain protein.
GN OrderedLocusNames=MCA1902;
OS Methylococcus capsulatus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
OX NCBI_TaxID=414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bath / NCIMB 11132;
RX PubMed=15383840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.L., Larsen O., Sakwa J., Bruseeth L., Khouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Methe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren Q., Read T.D., DeBoy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindhaug S.H., Holt I.E., Eidhammer I., Jonassen I., Vanaken S.,
RA Utterback T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R.,
RA Eisen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).";
RL PLOS Biol. 2:1616-1628(2004).
DR EMBL; AE017282; AAU91823.1; -; Genomic_DNA.
DR TIGR; MCA1902; -.
KW Complete proteome.
SQ SEQUENCE 178 AA; 19998 MW; 10916536F02CABC7 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 178;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
DB 111 GAVADEAAGW 120
```

```
RESULT 42
Q9HRC7_HALSA PRELIMINARY; PRT; 181 AA.
AC Q9HRC7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription initiation factor IIE alpha subunit.
GN Name=tfeA; OrderedLocusNames=VNG0757G;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welte R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005019; AAG19231.1; -; Genomic_DNA.
DR PIR; C84233; C84233.
DR GO; GO:0005673; C:transcription factor TFIIIE complex; IEA.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006367; P:transcription initiation from RNA polymeras. .; IEA.
DR InterPro; IPR002853; TFIIIE_alpha.
DR Pfam; PF02002; TFIIIE_alpha; 1.
KW Complete proteome; Initiation factor.
SQ SEQUENCE 181 AA; 20837 MW; 06185D1E43393B5D CRC64;

Query Match 66.1%; Score 37; DB 2; Length 181;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
DB 78 RVRDEDSGW 86

RESULT 43
ALBR_KLEOX STANDARD; PRT; 218 AA.
ID ALBR_KLEOX
AC P10488;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Albicidin resistance protein.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=JMP4505;
RX MEDLINE=89013885; PubMed=2845223;
RA Walker M.J., Birch R.G., Pemberton J.M.;
RT "Cloning and characterization of an albicidin resistance gene from
RT Klebsiella oxytoca.";
RL Mol. Microbiol. 2:443-454(1988).
CC -!- FUNCTION: Albicidin resistance protein binds to form a complex
CC without antibiotic activity but without catalyzing any further
CC chemical modifications to albicidin.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
-----
```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; Y00558; CAA68640.1; -; Genomic_DNA.
DR PIR; S02828; S02828.
DR InterPro; IPR008396; Albicidin_res.
DR Pfam; PF05583; Albicidin_res; 2.
KW Antibiotic resistance; Periplasmic.
SQ SEQUENCE 218 AA; 25859 MW; 56578223BEDC8CA7 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 218;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
||:|||||
Db 48 GRLRWQPAGW 57

RESULT 44

Q67JS4 SYMTH
ID Q67JS4_SYMTH PRELIMINARY; PRT; 244 AA.
AC Q67JS4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Formate dehydrogenase associated protein.
GN OrderedLocusNames=STH3094;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14863;
RX PubMed=15383646; DOI=10.1093/nar/gkh830;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable
bacterium that depends on microbial commensalism.";
RL Nucleic Acids Res. 32:4937-4944 (2004).
DR EMBL; AP006840; BAD42076.1; -; Genomic DNA.
DR GO; GO:0009326; C:formate dehydrogenase complex; IEA.
DR GO; GO:0008863; F:formate dehydrogenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003786; FdhD.
DR InterPro; IPR000408; Reg_chr_condens.
DR Pfam; PF02634; FdhD-NarQ; 1.
DR PROSITE; PS00626; RCC1_2; 1.
KW Complete proteome.
SQ SEQUENCE 244 AA; 25720 MW; F5566303CAE2ABD6 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 244;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RDQVAGW 10
|:|||||
Db 114 REQVAGW 120

RESULT 45

Q5UQ11 MIMIV
ID Q5UQ11_MIMIV PRELIMINARY; PRT; 276 AA.
AC Q5UQ11;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MIMI_L191;
OS Mimivirus.

OC Viruses; dsDNA viruses, no RNA stage; Mimivirus.
OX NCBI_TaxID=212035;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Rowbotham-Bradford;
RX MEDLINE=22550848; PubMed=12663918; DOI=10.1126/science.1081867;
RA La Scola B., Audic S., Robert C., Jungang L., de Lamballerie X.,
RA Drancourt M., Birtles R., Claverie J.M., Raoult D.;
RT "A giant virus in amoebae.";
RL Science 299:2033-2033 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Rowbotham-Bradford;
RX PubMed=15486256; DOI=10.1126/science.1101485;
RA Raoult D., Audic S., Robert C., Abergel C., Renesto P., Ogata H.,
RA La Scola B., Susan M., Claverie J.M.;
RT "The 1.2-Mb Genome Sequence of Mimivirus.";
RL Science 306:1344-1350 (2004).
DR EMBL; AY653733; AAV50465.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 276 AA; 32656 MW; 6EE1DBAC6433055C CRC64;

Query Match 66.1%; Score 37; DB 2; Length 276;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
||:|||||
Db 28 RVQDTIGW 36

RESULT 46

Q4UPY2_XANCP
ID Q4UPY2_XANCP PRELIMINARY; PRT; 307 AA.
AC Q4UPY2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=XC_3851;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000050; AAY50891.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 307 AA; 33115 MW; 141D08229A44B82A CRC64;

Query Match 66.1%; Score 37; DB 2; Length 307;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|:||||
Db 85 RIRQSIAGW 93

RESULT 47

Q4NPES_9DELT
ID Q4NPES_9DELT PRELIMINARY; PRT; 307 AA.
AC Q4NPES;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01001129; EAL42974.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 372 AA; 42436 MW; F5B8824993DD752D CRC64;

Query Match 66.1%; Score 37; DB 2; Length 372;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|:|:|
Db 111 GKVRDLIATW 120

Search completed: May 12, 2006, 10:48:54
Job time : 120.769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 72.6496 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-38
Perfect score: 49
Sequence: 1 ESVREQVMKY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	10	2	AAY22358 TPO recep
2	49	100.0	10	3	AAB16982 TPO-mimet
3	49	100.0	10	5	ABB72868 TPO mimet
4	49	100.0	10	7	ADJ73019 TPO mimet
5	49	100.0	10	8	ADJ52654 CH1 delet
6	49	100.0	10	8	ADJ51615 CH1 delet
7	38	77.6	291	6	ABU36234 Protein e
8	37	75.5	216	8	ADN99878 Novel hum
9	37	75.5	384	2	AAY20122 B. burgdo
10	37	75.5	429	2	AAY06993 Rat eukar
11	37	75.5	429	7	ADE60677 Rat Prote
12	37	75.5	429	7	ADD45400 Rat Prote
13	37	75.5	429	7	ADE60681 Rat Prote
14	37	75.5	429	7	ADE61162 Rat Prote
15	37	75.5	431	7	ADE61164 Human Pro
16	37	75.5	431	7	ADE60679 Human Pro
17	37	75.5	431	7	ADD45402 Human Pro
18	37	75.5	431	7	ADE60683 Human Pro
19	37	75.5	431	8	ADN04323 Antipsori
20	37	75.5	431	8	ABM81587 Tumour-as
21	37	75.5	431	8	ADP54723 Human PRO
22	37	75.5	431	9	ADX05561 Cyclin-de
23	37	75.5	431	9	ADY15998 PRO polyp
24	37	75.5	433	3	AAB43508 Human can

98	33	67.3	490	5	ABB55083	Abb55083 Lactococc	171	31	63.3	15	2	AAY22380	Aay22380 TPO recep
99	33	67.3	673	6	ABU33383	Abu33383 Protein e	172	31	63.3	55	4	AAM21674	Aam21674 Peptide #
100	33	67.3	673	9	AEb33948	Aeb33948 L. pneumo	173	31	63.3	55	4	AAM37981	Aam37981 Peptide #
101	33	67.3	687	9	AEb36554	Aeb36554 L. pneumo	174	31	63.3	55	4	AAM65049	Aam65049 Human bra
102	33	67.3	780	7	ABM88746	Abm88746 Rice abio	175	31	63.3	55	5	ABG46787	Abg46787 Human pep
103	33	67.3	793	5	ABJ17937	Abj17937 Rabbit ne	176	31	63.3	109	4	ABG19372	Abg19372 Novel hum
104	33	67.3	862	2	AAy07064	Aay07064 Renal can	177	31	63.3	109	4	ABG11683	Abg11683 Novel hum
105	33	67.3	862	4	AAB62199	Aab62199 Amino aci	178	31	63.3	120	6	ABP80598	Abp80598 N. gonorr
106	33	67.3	862	6	ABR47576	Abr47576 Breast ca	179	31	63.3	120	6	ABU37678	Abu37678 Protein e
107	33	67.3	862	7	ADJ70641	Adj70641 Human hea	180	31	63.3	120	6	ABU37281	Abu37281 Protein e
108	33	67.3	1153	6	ABU23672	Abu23672 Protein e	181	31	63.3	125	9	AEb38788	Aeb38788 L. pneumo
109	33	67.3	1530	7	ABM88229	Abm88229 Rice abio	182	31	63.3	131	6	AAE34014	Aae34014 Lolium pe
110	33	67.3	1873	1	AAP95645	Aap95645 Rabbit se	183	31	63.3	152	3	AAB44889	Aab44889 Human sec
111	33	67.3	1873	2	AAR73055	Aar73055 Rabbit sk	184	31	63.3	161	2	AAy35809	Aay35809 Amino aci
112	33	67.3	1873	2	AAW18390	Aaw18390 Rabbit ca	185	31	63.3	161	5	ABP28993	Abp28993 Streptoco
113	33	67.3	1873	2	AAW37711	Aaw37711 Rabbit sk	186	31	63.3	165	5	ABP28646	Abp28646 Streptoco
114	33	67.3	1873	3	AAy77544	Aay77544 Rabbit sk	187	31	63.3	165	8	ADV87871	Adv87871 Streptoco
115	33	67.3	1873	9	AEb26777	Aeb26777 Rabbit sk	188	31	63.3	165	8	ADV81323	Adv81323 Streptoco
116	32	65.3	14	2	AAy22396	Aay22396 TPO recep	189	31	63.3	165	8	ADV79124	Adv79124 Streptoco
117	32	65.3	15	2	AAy22399	Aay22399 TPO recep	190	31	63.3	165	9	AEb41756	Aeb41756 L. pneumo
118	32	65.3	15	2	AAy22381	Aay22381 TPO recep	191	31	63.3	175	9	ADx02798	Adx02798 Sendai vi
119	32	65.3	82	4	AAB63463	Aab63463 Human bre	192	31	63.3	176	9	AEb38502	Aeb38502 L. pneumo
120	32	65.3	125	5	AAE16594	Aae16594 Human 576	193	31	63.3	177	9	AEb38841	Aeb38841 L. pneumo
121	32	65.3	134	8	ADS18428	Ads18428 Entamoeba	194	31	63.3	181	9	ADx02797	Adx02797 Sendai vi
122	32	65.3	166	6	ABM69803	Abm69803 Photorhab	195	31	63.3	184	7	ADF03934	Adf03934 Bacterial
123	32	65.3	241	8	ADN26512	Adn26512 Bacterial	196	31	63.3	203	1	AAP70394	Aap70394 Sequence
124	32	65.3	251	8	ADQ36875	Adq36875 Cell prol	197	31	63.3	204	7	ADC85520	Adc85520 Sendai vi
125	32	65.3	275	6	ABU43401	Abu43401 Protein e	198	31	63.3	204	7	ADx02796	Adx02796 Sendai vi
126	32	65.3	315	7	ABM73684	Abm73684 DNA clone	199	31	63.3	206	6	ABU02574	Abu02574 S. pneumo
127	32	65.3	330	7	ABO65232	Abo65232 Klebsiell	200	31	63.3	206	8	ADK48267	Adk48267 Streptoco
128	32	65.3	349	4	AAE05984	Aae05984 Lactobaci	201	31	63.3	207	8	ADR94944	Adr94944 Novel S.
129	32	65.3	349	7	ADH89124	Adh89124 L. rhamno	202	31	63.3	207	9	AEA58814	Aea58814 Streptoco
130	32	65.3	376	8	ADN19201	Adn19201 Bacterial	203	31	63.3	214	7	ABO83235	Abo83235 Pseudomon
131	32	65.3	383	6	ABU25688	Abu25688 Protein e	204	31	63.3	215	8	ADx78070	Adx78070 Plant ful
132	32	65.3	398	7	ABM85352	Abm85352 Mouse pro	205	31	63.3	215	9	ADx02795	Adx02795 Sendai vi
133	32	65.3	402	8	ADS26395	Ads26395 Bacterial	206	31	63.3	226	8	ADK48371	Adk48371 Streptoco
134	32	65.3	403	8	ADS28482	Ads28482 Bacterial	207	31	63.3	246	8	ADR94895	Adr94895 Novel S.
135	32	65.3	403	8	ADN26837	Adn26837 Bacterial	208	31	63.3	246	9	AEA58765	Aea58765 Streptoco
136	32	65.3	403	8	ADS26763	Ads26763 Bacterial	209	31	63.3	258	6	ADA35565	Ada35565 Acinetoba
137	32	65.3	403	8	ADS27145	Ads27145 Bacterial	210	31	63.3	272	4	ABG11685	Abg11685 Novel hum
138	32	65.3	407	8	ADP84838	Adp84838 Cellobioh	211	31	63.3	279	3	AAB29865	Aab29865 Human sec
139	32	65.3	410	8	ADS44448	Ads44448 Bacterial	212	31	63.3	283	3	AAG13091	Aag13091 Arabidops
140	32	65.3	535	4	ABG15415	Abg15415 Novel hum	213	31	63.3	284	3	AAG13090	Aag13090 Arabidops
141	32	65.3	793	2	AAW25635	Aaw25635 Human cad	214	31	63.3	287	8	ADS22304	Ads22304 Bacterial
142	32	65.3	793	2	AAW13132	Aaw13132 Full leng	215	31	63.3	291	3	AAG13089	Aag13089 Arabidops
143	32	65.3	799	2	AAW25639	Aaw25639 Rat full	216	31	63.3	291	5	ABG91527	Abg91527 Purine/py
144	32	65.3	799	2	AAW13126	Aaw13126 Full leng	217	31	63.3	291	8	ADI43725	Adi43725 Plant tra
145	32	65.3	799	7	ADG42065	Adg42065 Human CDH	218	31	63.3	291	8	ADS43208	Ads43208 Bacterial
146	32	65.3	799	8	ADP26902	Adp26902 Human cad	219	31	63.3	291	9	AEA26845	Aea26845 Stress to
147	32	65.3	810	6	ABM70687	Abm70687 Photorhab	220	31	63.3	292	5	ABB48277	Abb48277 Listeria
148	32	65.3	824	4	ABB52696	Abb52696 Escherich	221	31	63.3	293	4	AAB96312	Aab96312 Putative
149	32	65.3	832	4	AAB84218	Aab84218 Amino aci	222	31	63.3	315	7	ADE57869	Ade57869 Human Pro
150	32	65.3	845	4	AAM80134	Aam80134 Human pro	223	31	63.3	315	7	ADE57877	Ade57877 Human Pro
151	32	65.3	847	6	ABM70680	Abm70680 Photorhab	224	31	63.3	315	7	ADE57881	Ade57881 Human Pro
152	32	65.3	848	7	ABO67256	Abo67256 Klebsiell	225	31	63.3	315	7	ADE57885	Ade57885 Human Pro
153	32	65.3	852	6	ABM70692	Abm70692 Photorhab	226	31	63.3	315	7	ADE57889	Ade57889 Human Pro
154	32	65.3	854	4	AAB84217	Aab84217 Amino aci	227	31	63.3	315	7	ADE57873	Ade57873 Human Pro
155	32	65.3	854	4	AAM79150	Aam79150 Human pro	228	31	63.3	315	8	ADP90915	Adp90915 Human ald
156	32	65.3	854	5	ABB81932	Abb81932 Human Dis	229	31	63.3	315	9	ADV85532	Adv85532 Human ald
157	32	65.3	854	8	ADK17424	Adk17424 Human dis	230	31	63.3	316	2	AAR06652	Aar06652 Placenta-
158	32	65.3	854	8	ADR47379	Adr47379 Human DIS	231	31	63.3	316	2	AAR15425	Aar15425 Human ald
159	32	65.3	863	8	ADR47391	Adr47391 DISC1-HA.	232	31	63.3	316	5	ABB04753	Abb04753 Human AKR
160	32	65.3	896	4	AAM00820	Aam00820 Human bon	233	31	63.3	316	5	ABB77443	Abb77443 Human tum
161	32	65.3	940	6	ABU49109	Abu49109 Protein e	234	31	63.3	316	7	ADJ69767	Adj69767 Human can
162	32	65.3	1195	8	ADS21995	Ads21995 Bacterial	235	31	63.3	316	8	ADQ15140	Adq15140 Human can
163	32	65.3	1231	7	ADJ68516	Adj68516 Human hea	236	31	63.3	316	8	ABM80940	Abm80940 Tumour-as
164	32	65.3	1246	8	ADQ97435	Adq97435 Human can	237	31	63.3	316	8	ADP23164	Adp23164 PRO polyp
165	32	65.3	1246	8	ABM80794	Abm80794 Tumour-as	238	31	63.3	316	9	AEB22075	Aeb22075 Human AKR
166	32	65.3	1266	8	ADO51507	Ado51507 Human jum	239	31	63.3	318	5	AAE15443	Aae15443 Human dru
167	32	65.3	1456	7	ADM04577	Adm04577 Human pro	240	31	63.3	326	9	AEB42007	Aeb42007 L. pneumo
168	32	65.3	1620	6	ABP98333	Abp98333 Amino aci	241	31	63.3	328	4	ABG11687	Abg11687 Novel hum
169	32	65.3	2434	3	AAB42264	Aab42264 Human ORF	242	31	63.3	347	9	AEB42011	Aeb42011 L. pneumo
170	32	65.3	2594	4	AAM00984	Aam00984 Human bon	243	31	63.3	348	6	ABU17065	Abu17065 Protein e

244	31	63.3	356	6	ADA33310	Ada33310 Acinetoba	317	30	61.2	223	7	ADC10983	Adc10983 Human pro
245	31	63.3	391	9	AEb42009	Aeb42009 L. pneumo	318	30	61.2	232	9	AEb98137	Aeb98137 Protein e
246	31	63.3	391	9	AEb41974	Aeb41974 L. pneumo	319	30	61.2	238	4	AAG91726	Aag91726 C glutami
247	31	63.3	391	9	AEb41956	Aeb41956 L. pneumo	320	30	61.2	238	6	ABR44270	Abr44270 Novel hum
248	31	63.3	391	9	AEb42013	Aeb42013 L. pneumo	321	30	61.2	238	6	ABR44269	Abr44269 Novel hum
249	31	63.3	394	4	ABG19373	Abg19373 Novel hum	322	30	61.2	238	7	ADE14367	Ade14367 Human int
250	31	63.3	402	9	AEb38745	Aeb38745 L. pneumo	323	30	61.2	238	8	ABO84778	Abo84778 Human can
251	31	63.3	402	9	AEb38771	Aeb38771 L. pneumo	324	30	61.2	238	8	ADN25862	Adn25862 Bacterial
252	31	63.3	416	6	AAE34863	Aae34863 Human kin	325	30	61.2	241	4	AAG98328	Aag98328 Escherich
253	31	63.3	416	8	ADK65919	Adk65919 Human IPM	326	30	61.2	249	8	ADS28386	Ads28386 Bacterial
254	31	63.3	433	8	ADK16215	Adk16215 Nanoarcha	327	30	61.2	251	9	ADR95327	Adr95327 Novel S.
255	31	63.3	437	6	ABU33339	Abu33339 Protein e	328	30	61.2	251	9	AEA59197	Aea59197 Streptoco
256	31	63.3	437	9	AEb36769	Aeb36769 L. pneumo	329	30	61.2	254	3	AAG52497	Aag52497 Arabidops
257	31	63.3	437	9	AEb40152	Aeb40152 L. pneumo	330	30	61.2	255	3	AAG52496	Aag52496 Arabidops
258	31	63.3	450	6	ADA55062	Ada55062 Human pro	331	30	61.2	256	8	ABO59795	Abo59795 Human gen
259	31	63.3	451	5	ABP25826	Abp25826 Streptoco	332	30	61.2	283	8	ADN46138	Adn46138 Thermococ
260	31	63.3	451	8	ADV89528	Adv89528 Streptoco	333	30	61.2	285	2	AAy35437	Aay35437 Chlamydia
261	31	63.3	451	8	ADV80781	Adv80781 Streptoco	334	30	61.2	286	9	ABM96030	Abm96030 M. xanthu
262	31	63.3	451	8	ADV82942	Adv82942 Streptoco	335	30	61.2	291	7	ABM85353	Abm85353 Human pro
263	31	63.3	459	4	ABG11686	Abg11686 Novel hum	336	30	61.2	292	3	AAG52495	Aag52495 Arabidops
264	31	63.3	496	6	ABU33143	Abu33143 Protein e	337	30	61.2	295	5	ABG91517	Abg91517 Purine/py
265	31	63.3	496	9	AEb40962	Aeb40962 L. pneumo	338	30	61.2	302	4	AAb94000	Aab94000 Human pro
266	31	63.3	496	9	AEb37651	Aeb37651 L. pneumo	339	30	61.2	308	7	ABO79435	Abo79435 Pseudomon
267	31	63.3	637	8	ADX66365	Adx66365 Plant ful	340	30	61.2	315	8	ADN47480	Adn47480 Thermococ
268	31	63.3	644	6	ABU25152	Abu25152 Protein e	341	30	61.2	332	8	ABO84777	Abo84777 Murine ca
269	31	63.3	687	2	AAR97123	Aar97123 Chicken d	342	30	61.2	336	7	ABO65485	Abo65485 Klebsiell
270	31	63.3	729	8	ADN27264	Adn27264 Bacterial	343	30	61.2	339	9	ADZ25527	Adz25527 Cytokine
271	31	63.3	737	8	ADK47776	Adk47776 Streptoco	344	30	61.2	344	4	ABG26158	Abg26158 Novel hum
272	31	63.3	744	8	ADR95419	Adr95419 Novel S.	345	30	61.2	346	8	ADN22477	Adn22477 Bacterial
273	31	63.3	744	9	AEA59289	Aea59289 Streptoco	346	30	61.2	347	3	AAG51011	Aag51011 Arabidops
274	31	63.3	769	8	ADS27589	Ads27589 Bacterial	347	30	61.2	348	4	AAb76604	Aab76604 Corynebac
275	31	63.3	782	5	AAU98072	Aau98072 Chicken n	348	30	61.2	348	4	AAb79964	Aab79964 Corynebac
276	31	63.3	794	6	ABU17842	Abul17842 Protein e	349	30	61.2	348	4	AAG92166	Aag92166 C glutami
277	31	63.3	842	4	AAG85019	Aag85019 Shrimp wh	350	30	61.2	350	5	AAO17160	Aao17160 Corynebac
278	31	63.3	863	4	ABB58669	Abb58669 Drosophil	351	30	61.2	350	4	AAb48049	Aab48049 Signal tr
279	31	63.3	924	5	ABB54116	Abb54116 Lactococc	352	30	61.2	351	4	AAb62149	Aab62149 P. falcip
280	31	63.3	924	8	ADS29350	Ads29350 Bacterial	353	30	61.2	355	3	AAG28692	Aag28692 Arabidops
281	31	63.3	932	6	ABU19329	Abul19329 Protein e	354	30	61.2	359	4	AAb30608	Aab30608 Amino aci
282	31	63.3	978	8	ADQ97967	Adq97967 Mouse can	355	30	61.2	359	5	ABP65172	Abp65172 Hypoxia-r
283	31	63.3	990	9	ABM92566	Abm92566 M. xanthu	356	30	61.2	359	8	ADP54847	Adp54847 Human PRO
284	31	63.3	1016	8	ADR08525	Adr08525 Human pro	357	30	61.2	380	8	ADQ19952	Adq19952 Human sof
285	31	63.3	1164	8	ADN23911	Adn23911 Bacterial	358	30	61.2	382	6	ABU33633	Abu33633 Protein e
286	31	63.3	1215	4	AAE02340	Aae02340 D. melano	359	30	61.2	382	8	ADS27889	Ads27889 Bacterial
287	31	63.3	1264	8	ADQ67800	Adq67800 Novel hum	360	30	61.2	382	9	AEb39551	Aeb39551 L. pneumo
288	31	63.3	2172	4	ABB67958	Abb67958 Drosophil	361	30	61.2	385	4	ABB58316	Abb58316 Drosophil
289	31	63.3	2324	9	ABM90533	Abm90533 M. xanthu	362	30	61.2	392	3	AAG51010	Aag51010 Arabidops
290	30	61.2	11	2	AAy22367	Aay22367 TPO recep	363	30	61.2	394	6	ADB10152	Adb10152 Alloiococ
291	30	61.2	15	2	AAy22400	Aay22400 TPO recep	364	30	61.2	399	3	AAG28691	Aag28691 Arabidops
292	30	61.2	51	5	ABP34908	Abp34908 Human ORF	365	30	61.2	400	3	AAg85362	Aag85362 Human pho
293	30	61.2	62	5	ABP04777	Abp04777 Human ORF	366	30	61.2	402	4	AAb85362	Aab85362 Human pho
294	30	61.2	93	9	AEb41682	Aeb41682 L. pneumo	367	30	61.2	411	5	ABP69581	Abp69581 Human pol
295	30	61.2	94	4	AAU01705	Aau01705 Propeptid	368	30	61.2	422	8	ADQ65892	Adq65892 Novel hum
296	30	61.2	99	9	AEb38411	Aeb38411 L. pneumo	369	30	61.2	430	5	ABB55320	Abb55320 Lactococc
297	30	61.2	118	6	ABU50358	Abu50358 Protein e	370	30	61.2	432	6	ABM73527	Abm73527 Staphyloc
298	30	61.2	150	7	ADC89106	Adc89106 Ribosomal	371	30	61.2	434	4	AAm25227	Aam25227 Human pro
299	30	61.2	171	3	AAb58741	Aab58741 Breast an	372	30	61.2	445	3	AAG51009	Aag51009 Arabidops
300	30	61.2	171	3	ABP41060	Abp41060 Human ova	373	30	61.2	446	6	ABU39416	Abu39416 Protein e
301	30	61.2	171	5	AAE33498	Aae33498 Streptoco	374	30	61.2	448	8	ADX94435	Adx94435 Plant ful
302	30	61.2	171	6	AAE33498	Aae33498 Streptoco	375	30	61.2	453	3	AAG28690	Aag28690 Arabidops
303	30	61.2	176	4	AAU39170	Aau39170 Propionib	376	30	61.2	461	6	ABU19541	Abul19541 Protein e
304	30	61.2	176	5	ABP10173	Abp10173 Human ORF	377	30	61.2	466	4	AAb96071	Aab96071 Putative
305	30	61.2	176	6	ABM35689	Abm35689 Propionib	378	30	61.2	466	8	ADN18586	Adn18586 Bacterial
306	30	61.2	177	3	AAb51685	Aab51685 Gene 22 h	379	30	61.2	476	7	ADJ71050	Adj71050 Human hea
307	30	61.2	177	3	AAb51686	Aab51686 Human sec	380	30	61.2	487	4	AAU19641	Aau19641 Human nov
308	30	61.2	177	8	ABO59208	Abo59208 Human gen	381	30	61.2	487	5	ABP47861	Abp47861 Human pol
309	30	61.2	185	3	AAG58376	Aag58376 Arabidops	382	30	61.2	487	6	ABU17480	Abul17480 Protein e
310	30	61.2	202	9	AEb98133	Aeb98133 Protein e	383	30	61.2	487	7	ADC10823	Adc10823 Human ext
311	30	61.2	205	8	ADS24392	Ads24392 Bacterial	384	30	61.2	489	8	ADS29678	Ads29678 Bacterial
312	30	61.2	210	3	AAG58375	Aag58375 Arabidops	385	30	61.2	517	8	ADI35156	Adi35156 C. tentan
313	30	61.2	210	8	ADT56711	Adt56711 Plant pol	386	30	61.2	538	8	ADS29146	Ads29146 Bacterial
314	30	61.2	212	9	AEb98131	Aeb98131 Protein e	387	30	61.2	546	2	AAW81975	Aaw81975 Ehrlichia
315	30	61.2	223	4	AAU19801	Aau19801 Human nov	388	30	61.2	551	7	ABO67710	Abo67710 Klebsiell
316	30	61.2	223	5	ABP48021	Abp48021 Human pol	389	30	61.2	582	4	ABB69962	Abb69962 Drosophil

390 30 61.2 583 8 ADT58999 Adt58999 Plant pol 463
391 30 61.2 613 4 AAM93828 Aam93828 Human pol 464
392 30 61.2 613 8 ADL31858 Adl31858 Human pro 465
393 30 61.2 615 4 ABG26157 Abg26157 Novel hum 466
394 30 61.2 620 4 ABG25813 Abg25813 Novel hum 467
395 30 61.2 669 4 ABB59622 Abb59622 Drosophil 468
396 30 61.2 681 9 ADY70275 Ady70275 Human bet 469
397 30 61.2 688 8 ADU06846 Adu06846 Cell adhe 470
398 30 61.2 694 7 ADB85208 Adb85208 Rat cadhe 471
399 30 61.2 694 7 ADD46851 Add46851 Rat Prote 472
400 30 61.2 698 6 ABU16049 Abul6049 Protein e 473
401 30 61.2 698 6 ABM72008 Abm72008 Staphyloc 474
402 30 61.2 705 4 AAU36870 Aau36870 Staphyloc 475
403 30 61.2 727 8 ADN22992 Adn22992 Bacterial 476
404 30 61.2 731 8 ADQ66540 Adq66540 Novel hum 477
405 30 61.2 736 8 ADR86001 Adr86001 Aspergill 478
406 30 61.2 744 6 ABU37141 Abu37141 Protein e 479
407 30 61.2 747 7 ADD14035 Add14035 Human src 480
408 30 61.2 759 4 ABG21947 Abg21947 Novel hum 481
409 30 61.2 759 6 ABP78828 Abp78828 N. gonorr 482
410 30 61.2 759 6 ABU38003 Abu38003 Protein e 483
411 30 61.2 761 8 ADP29437 Adp29437 Human sec 484
412 30 61.2 769 1 AAP90954 Aap90954 Yeast top 485
413 30 61.2 769 6 ABR53198 Abr53198 Protein s 486
414 30 61.2 769 7 ADK63244 Adk63244 Disease t 487
415 30 61.2 769 8 ADS43918 Ads43918 Bacterial 488
416 30 61.2 771 8 ADN21684 Adn21684 Bacterial 489
417 30 61.2 771 8 ADN24439 Adn24439 Bacterial 490
418 30 61.2 772 2 AAY41725 Aay41725 Human PRO 491
419 30 61.2 772 3 AAB44281 Aab44281 Human PRO 492
420 30 61.2 772 4 AAB31209 Aab31209 Amino aci 493
421 30 61.2 772 4 AAU29076 Aau29076 Human PRO 494
422 30 61.2 772 6 ABU58452 Abu58452 Human PRO 495
423 30 61.2 772 6 ABU88000 Abu88000 Novel hum 496
424 30 61.2 772 6 ABU84315 Abu84315 Human sec 497
425 30 61.2 772 6 ABR66189 Abr66189 Human sec 498
426 30 61.2 772 6 ABR65579 Abr65579 Human sec 499
427 30 61.2 772 6 ABU99519 Abu99519 Human sec 500
428 30 61.2 772 6 ABU82758 Abu82758 Human PRO 501
429 30 61.2 772 6 ABU89879 Abu89879 Novel hum 502
430 30 61.2 772 6 ABR68128 Abr68128 Human sec 503
431 30 61.2 772 6 ABU96181 Abu96181 Novel hum 504
432 30 61.2 772 6 ABU92612 Abu92612 Human sec 505
433 30 61.2 772 6 ABO08689 Abo08689 Human sec 506
434 30 61.2 772 6 ABO02741 Abo02741 Human sec 507
435 30 61.2 772 6 ABR74895 Abr74895 Human sec 508
436 30 61.2 772 6 ABR94657 Abr94657 Human sec 509
437 30 61.2 772 6 ABO25180 Abo25180 Novel hum 510
438 30 61.2 772 6 ABO25227 Abo25227 Novel hum 511
439 30 61.2 772 6 ABU85630 Abu85630 Human PRO 512
440 30 61.2 772 6 ABU98790 Abu98790 Novel hum 513
441 30 61.2 772 6 ABU98005 Abu98005 Novel hum 514
442 30 61.2 772 6 ABU91711 Abu91711 Novel hum 515
443 30 61.2 772 6 ABU72233 Abu72233 Novel hum 516
444 30 61.2 772 6 ABU89404 Abu89404 Human PRO 517
445 30 61.2 772 6 ABU86245 Abu86245 Human sec 518
446 30 61.2 772 6 ABU67458 Abu67458 Human sec 519
447 30 61.2 772 6 ABU80486 Abu80486 Human PRO 520
448 30 61.2 772 6 ABR99404 Abr99404 Human sec 521
449 30 61.2 772 6 ABR98794 Abr98794 Human sec 522
450 30 61.2 772 6 ABO16317 Abo16317 Human sec 523
451 30 61.2 772 6 ABR92217 Abr92217 Human sec 524
452 30 61.2 772 6 ABO18858 Abo18858 Human sec 525
453 30 61.2 772 6 ABR78279 Abr78279 Human sec 526
454 30 61.2 772 6 ABU85015 Abu85015 Novel hum 527
455 30 61.2 772 6 ABO00154 Abo00154 Novel hum 528
456 30 61.2 772 6 ABO11486 Abo11486 Human sec 529
457 30 61.2 772 6 ABO02131 Abo02131 Human sec 530
458 30 61.2 772 6 ABU88705 Abu88705 Novel hum 531
459 30 61.2 772 6 ABU67298 Abu67298 Novel hum 532
460 30 61.2 772 6 ABU83400 Abu83400 Human sec 533
461 30 61.2 772 6 ABO06201 Abo06201 Novel hum 534
462 30 61.2 772 6 ABR59237 Abr59237 Human sec 535

772 6 ABO09299 Abo09299 Human sec
772 6 ABO19163 Abo19163 Novel hum
772 6 ABO11181 Abo11181 Human sec
772 6 ABR66799 Abr66799 Human sec
772 6 ABO16012 Abo16012 Human sec
772 6 ABO13718 Abo13718 Human sec
772 6 ABU84913 Abu84913 Human sec
772 6 ABU65621 Abu65621 Human sec
772 6 ABO07469 Abo07469 Human PRO
772 6 ABO03656 Abo03656 Human sec
772 6 ABR67104 Abr67104 Human sec
772 6 ABO15707 Abo15707 Human sec
772 6 ABU55988 Abu55988 Human sec
772 6 ABU61111 Abu61111 Human PRO
772 6 ABU65316 Abu65316 Human PRO
772 6 ABU95261 Abu95261 Novel hum
772 6 ABU71164 Abu71164 Human PRO
772 6 ABO07774 Abo07774 Human PRO
772 6 ABR70015 Abr70015 Human sec
772 6 ABR69348 Abr69348 Human sec
772 6 ABO01489 Abo01489 Human PRO
772 6 ABU81291 Abu81291 Human PRO
772 6 ABR60088 Abr60088 Human sec
772 6 ABR67823 Abr67823 Human sec
772 6 ABR65211 Abr65211 Human sec
772 6 ABR68433 Abr68433 Human sec
772 6 ABR71845 Abr71845 Human sec
772 6 ABU85325 Abu85325 Human PRO
772 6 ABU89015 Abu89015 Human sec
772 6 ABU83095 Abu83095 Human sec
772 6 ABU94951 Abu94951 Novel hum
772 6 ABU90499 Abu90499 Novel hum
772 6 ABU84010 Abu84010 Human sec
772 6 ABU93661 Abu93661 Novel hum
772 6 ABR64906 Abr64906 Human sec
772 6 ABR68738 Abr68738 Human sec
772 6 ABO06554 Abo06554 Human sec
772 6 ABR99099 Abr99099 Human sec
772 6 ABU56983 Abu56983 Human PRO
772 6 ABU72066 Abu72066 Novel hum
772 6 ABU85935 Abu85935 Novel hum
772 6 ABU82222 Abu82222 Novel hum
772 6 ABU87233 Abu87233 Human PRO
772 6 ABU67167 Abu67167 Novel hum
772 6 ABU83705 Abu83705 Human sec
772 6 ABO08079 Abo08079 Human PRO
772 6 ABU81790 Abu81790 Novel hum
772 6 ABU65954 Abu65954 Novel hum
772 6 ABR59783 Abr59783 Human sec
772 6 ABU93971 Abu93971 Novel hum
772 6 ABU80380 Abu80380 Human sec
772 6 ABU99824 Abu99824 Novel hum
772 6 ABR66494 Abr66494 Human sec
772 6 ABR90912 Abr90912 Human sec
772 6 ABU94339 Abu94339 Human PRO
772 6 ABU79221 Abu79221 Human PRO
772 6 ABU86550 Abu86550 Human sec
772 6 ABU86855 Abu86855 Novel hum
772 6 ABU94644 Abu94644 Human PRO
772 6 ABO04571 Abo04571 Human PRO
772 6 ABR70320 Abr70320 Human sec
772 6 ABU98485 Abu98485 Human PRO
772 6 ABR65884 Abr65884 Human sec
772 6 ABR64601 Abr64601 Human sec
772 6 ABU79526 Abu79526 Human PRO
772 6 ABU92917 Abu92917 Human sec
772 6 ABU95876 Abu95876 Human PRO
772 6 ABU91096 Abu91096 Novel hum
772 6 ABU90189 Abu90189 Novel hum
772 6 ABO09604 Abo09604 Human sec
772 6 ABO10876 Abo10876 Human sec
772 6 ABR70930 Abr70930 Human sec
772 6 ABU87538 Abu87538 Human PRO

536	30	61.2	772	6	ABU91406	Abu91406 Human PRO	609	30	61.2	772	6	ABO27862	Abo27862 Human sec
537	30	61.2	772	6	ABU84620	Abu84620 Human sec	610	30	61.2	772	6	ABO29997	Abo29997 Human sec
538	30	61.2	772	6	ABR69710	AbR69710 Human sec	611	30	61.2	772	6	ABO33206	Abo33206 Human PRO
539	30	61.2	772	6	ABU80087	Abu80087 Human PRO	612	30	61.2	772	6	ABM04894	Abm04894 Human sec
540	30	61.2	772	6	ABU93356	Abu93356 Human PRO	613	30	61.2	772	6	ABM08854	Abm08854 Human sec
541	30	61.2	772	6	ABO09909	Abo09909 Human sec	614	30	61.2	772	6	ABO36454	Abo36454 Human sec
542	30	61.2	772	6	ABO08994	Abo08994 Human sec	615	30	61.2	772	6	ABO35539	Abo35539 Human PRO
543	30	61.2	772	6	ABU10562	Abu10562 Human sec	616	30	61.2	772	6	ABO39504	Abo39504 Human sec
544	30	61.2	772	6	ABU95571	Abu95571 Human PRO	617	30	61.2	772	6	ABM10379	Abm10379 Human sec
545	30	61.2	772	6	ABU79809	Abu79809 Human sec	618	30	61.2	772	6	ABM11904	Abm11904 Human sec
546	30	61.2	772	6	ABU96780	Abu96780 Novel hum	619	30	61.2	772	6	ABO52050	Abo52050 Human PRO
547	30	61.2	772	6	ABR70625	AbR70625 Human sec	620	30	61.2	772	6	ABO52355	Abo52355 Human PRO
548	30	61.2	772	6	ABO04976	Abo04976 Novel hum	621	30	61.2	772	6	ABO23673	Abo23673 Human sec
549	30	61.2	772	6	ABO08384	Abo08384 Human sec	622	30	61.2	772	6	ABR97159	AbR97159 Human sec
550	30	61.2	772	6	ABO05591	Abo05591 Human sec	623	30	61.2	772	6	ABR86947	AbR86947 Human sec
551	30	61.2	772	6	ABR73980	AbR73980 Human sec	624	30	61.2	772	6	ABM10989	Abm10989 Human sec
552	30	61.2	772	6	ABR95572	AbR95572 Human sec	625	30	61.2	772	6	ABM28133	Abm28133 Human sec
553	30	61.2	772	6	ABR80869	AbR80869 Human sec	626	30	61.2	772	6	ABO32132	Abo32132 Human sec
554	30	61.2	772	6	ABR81174	AbR81174 Human sec	627	30	61.2	772	6	ABM15259	Abm15259 Human sec
555	30	61.2	772	6	ABM00870	Abm00870 Human sec	628	30	61.2	772	6	ABM06414	Abm06414 Human sec
556	30	61.2	772	6	ABR88472	AbR88472 Human sec	629	30	61.2	772	6	ABM04225	Abm04225 Human sec
557	30	61.2	772	6	ABM77293	Abm77293 Human sec	630	30	61.2	772	6	ABM22338	Abm22338 Human sec
558	30	61.2	772	6	ABO28777	Abo28777 Human sec	631	30	61.2	772	6	ABM07634	Abm07634 Human sec
559	30	61.2	772	6	ABO31522	Abo31522 Human sec	632	30	61.2	772	6	ABO40724	Abo40724 Human sec
560	30	61.2	772	6	ABM07939	Abm07939 Human sec	633	30	61.2	772	6	ABM35371	Abm35371 Human sec
561	30	61.2	772	6	ABO40419	Abo40419 Human sec	634	30	61.2	772	6	ABM33134	Abm33134 Human sec
562	30	61.2	772	6	ABO35844	Abo35844 Human PRO	635	30	61.2	772	6	ABO52660	Abo52660 Human PRO
563	30	61.2	772	6	ABO43983	Abo43983 Human PRO	636	30	61.2	772	6	ABO50220	Abo50220 Human sec
564	30	61.2	772	6	ADA77858	Ada77858 Human sec	637	30	61.2	772	6	ABU99214	Abu99214 Human sec
565	30	61.2	772	6	ABM24778	Abm24778 Human sec	638	30	61.2	772	6	ABO04266	Abo04266 Human sec
566	30	61.2	772	6	ABO03046	Abo03046 Human sec	639	30	61.2	772	6	ABO05896	Abo05896 Human sec
567	30	61.2	772	6	ABR90302	AbR90302 Human sec	640	30	61.2	772	6	ABM18436	Abm18436 Human sec
568	30	61.2	772	6	ABM17216	Abm17216 Human sec	641	30	61.2	772	6	ABR97464	AbR97464 Human sec
569	30	61.2	772	6	ABR94962	AbR94962 Human sec	642	30	61.2	772	6	ABR80564	AbR80564 Human sec
570	30	61.2	772	6	ABR95267	AbR95267 Human sec	643	30	61.2	772	6	ABM01175	Abm01175 Human sec
571	30	61.2	772	6	ABO21505	Abo21505 Human sec	644	30	61.2	772	6	ABR88777	AbR88777 Human sec
572	30	61.2	772	6	ABR97769	AbR97769 Human sec	645	30	61.2	772	6	ABM13429	Abm13429 Human sec
573	30	61.2	772	6	ABR87557	AbR87557 Human sec	646	30	61.2	772	6	ABM20813	Abm20813 Human sec
574	30	61.2	772	6	ABM77598	Abm77598 Human sec	647	30	61.2	772	6	ABO41944	Abo41944 Human sec
575	30	61.2	772	6	ABM27828	Abm27828 Human sec	648	30	61.2	772	6	ABO42554	Abo42554 Human sec
576	30	61.2	772	6	ABM06109	Abm06109 Human sec	649	30	61.2	772	6	ABM10074	Abm10074 Human sec
577	30	61.2	772	6	ABM03615	Abm03615 Human sec	650	30	61.2	772	6	ABO38589	Abo38589 Human sec
578	30	61.2	772	6	ABM35066	Abm35066 Human sec	651	30	61.2	772	6	ABM32829	Abm32829 Human sec
579	30	61.2	772	6	ABM26303	Abm26303 Human sec	652	30	61.2	772	6	ABM22643	Abm22643 Human sec
580	30	61.2	772	6	ABO48085	Abo48085 Human sec	653	30	61.2	772	6	ABM74854	Abm74854 Human sec
581	30	61.2	772	6	ABR92827	AbR92827 Human sec	654	30	61.2	772	6	ADA79650	Ada79650 Human sec
582	30	61.2	772	6	ABO24588	Abo24588 Human sec	655	30	61.2	772	6	ADA24803	Ada24803 Novel hum
583	30	61.2	772	6	ABM11599	Abm11599 Human sec	656	30	61.2	772	6	ABR96244	AbR96244 Human sec
584	30	61.2	772	6	ABM02700	Abm02700 Human sec	657	30	61.2	772	6	ABM02395	Abm02395 Human sec
585	30	61.2	772	6	ABM15996	Abm15996 Human sec	658	30	61.2	772	6	ABR86337	AbR86337 Human sec
586	30	61.2	772	6	ABO27557	Abo27557 Human sec	659	30	61.2	772	6	ABR86642	AbR86642 Human sec
587	30	61.2	772	6	ABM29048	Abm29048 Human sec	660	30	61.2	772	6	ABM16606	Abm16606 Human sec
588	30	61.2	772	6	ABM07024	Abm07024 Human sec	661	30	61.2	772	6	ABM29658	Abm29658 Human sec
589	30	61.2	772	6	ABM21118	Abm21118 Human sec	662	30	61.2	772	6	ABO29082	Abo29082 Human sec
590	30	61.2	772	6	ABM09464	Abm09464 Human sec	663	30	61.2	772	6	ABM23863	Abm23863 Human sec
591	30	61.2	772	6	ABO41334	Abo41334 Human sec	664	30	61.2	772	6	ABM23253	Abm23253 Human sec
592	30	61.2	772	6	ABO36149	Abo36149 Human PRO	665	30	61.2	772	6	ADA47293	Ada47293 Human sec
593	30	61.2	772	6	ABO43678	Abo43678 Human sec	666	30	61.2	772	6	ABM22033	Abm22033 Human sec
594	30	61.2	772	6	ABM76378	Abm76378 Human sec	667	30	61.2	772	6	ABO37674	Abo37674 Human sec
595	30	61.2	772	6	ABM76074	Abm76074 Human sec	668	30	61.2	772	6	ABM28438	Abm28438 Human sec
596	30	61.2	772	6	ABM25693	Abm25693 Human sec	669	30	61.2	772	6	ABM28743	Abm28743 Human sec
597	30	61.2	772	6	ABM25998	Abm25998 Human sec	670	30	61.2	772	6	ABM66387	Abm66387 Human sec
598	30	61.2	772	6	ABO03351	Abo03351 Human sec	671	30	61.2	772	6	ABM75769	Abm75769 Human sec
599	30	61.2	772	6	ABO02436	Abo02436 Human sec	672	30	61.2	772	6	ABM34049	Abm34049 Human sec
600	30	61.2	772	6	ABR90607	AbR90607 Human sec	673	30	61.2	772	6	ABM34354	Abm34354 Human sec
601	30	61.2	772	6	ABR73675	AbR73675 Human sec	674	30	61.2	772	6	ABO19682	Abo19682 Novel hum
602	30	61.2	772	6	ABO16927	Abo16927 Human sec	675	30	61.2	772	6	ABO20285	Abo20285 Human sec
603	30	61.2	772	6	ABR94352	AbR94352 Human sec	676	30	61.2	772	6	ABO21200	Abo21200 Human sec
604	30	61.2	772	6	ABR75859	AbR75859 Human sec	677	30	61.2	772	6	ABO22115	Abo22115 Human sec
605	30	61.2	772	6	ABR71235	AbR71235 Human sec	678	30	61.2	772	6	ADA12464	Ada12464 Human sec
606	30	61.2	772	6	ABR93132	AbR93132 Human sec	679	30	61.2	772	6	ABR96549	AbR96549 Human sec
607	30	61.2	772	6	ABR93437	AbR93437 Human sec	680	30	61.2	772	6	ABR85727	AbR85727 Human sec
608	30	61.2	772	6	ABR87862	AbR87862 Human sec	681	30	61.2	772	6	ABR99709	AbR99709 Human sec

682	30	61.2	772	6	ABM00565	Abm00565 Human sec	755	30	61.2	772	6	ABM15564	Abm15564 Human sec
683	30	61.2	772	6	ABM00260	Abm00260 Human sec	756	30	61.2	772	6	ABM08549	Abm08549 Human sec
684	30	61.2	772	6	ABO29692	Abo29692 Human sec	757	30	61.2	772	6	ABO42249	Abo42249 Human sec
685	30	61.2	772	6	ABM23558	Abm23558 Human sec	758	30	61.2	772	6	ABO37979	Abo37979 Human sec
686	30	61.2	772	6	ABM29353	Abm29353 Human sec	759	30	61.2	772	6	ABO45889	Abo45889 Human PRO
687	30	61.2	772	6	ABO38284	Abo38284 Human sec	760	30	61.2	772	6	ABM66692	Abm66692 Human sec
688	30	61.2	772	6	ABO45584	Abo45584 Human PRO	761	30	61.2	772	6	ADB20218	Adb20218 Human sec
689	30	61.2	772	6	ABM20508	Abm20508 Human sec	762	30	61.2	772	6	ABM19593	Abm19593 Human sec
690	30	61.2	772	6	ADA81377	Ada81377 Human sec	763	30	61.2	772	6	ABO49305	Abo49305 Human sec
691	30	61.2	772	6	ABO16622	Abo16622 Human sec	764	30	61.2	772	6	ABO49610	Abo49610 Human sec
692	30	61.2	772	6	ABO18248	Abo18248 Human sec	765	30	61.2	772	6	ADA78470	Ada78470 Human sec
693	30	61.2	772	6	ABO22675	Abo22675 Human PRO	766	30	61.2	772	6	ABO19573	Abo19573 Novel hum
694	30	61.2	772	6	ABO22980	Abo22980 Human PRO	767	30	61.2	772	6	ABR88167	Abr88167 Human sec
695	30	61.2	772	6	ABR92522	Abr92522 Human sec	768	30	61.2	772	6	ABM26913	Abm26913 Human sec
696	30	61.2	772	6	ABR81479	Abr81479 Human sec	769	30	61.2	772	6	ABM03310	Abm03310 Human sec
697	30	61.2	772	6	ABM77903	Abm77903 Human sec	770	30	61.2	772	7	ABO39809	Abo39809 Human sec
698	30	61.2	772	6	ABR89692	Abr89692 Human sec	771	30	61.2	772	7	ABO49915	Abo49915 Human sec
699	30	61.2	772	6	ABM26608	Abm26608 Human sec	772	30	61.2	772	7	ABO50830	Abo50830 Human sec
700	30	61.2	772	6	ABM13734	Abm13734 Human sec	773	30	61.2	772	7	ABO05286	Abo05286 Human sec
701	30	61.2	772	6	ABO28472	Abo28472 Human sec	774	30	61.2	772	7	ABR74590	Abr74590 Human sec
702	30	61.2	772	6	ABO30302	Abo30302 Human sec	775	30	61.2	772	7	ABR77069	Abr77069 Human sec
703	30	61.2	772	6	ABM07329	Abm07329 Human sec	776	30	61.2	772	7	ABM17826	Abm17826 Human sec
704	30	61.2	772	6	ABM03920	Abm03920 Human sec	777	30	61.2	772	7	ABR95877	Abr95877 Human sec
705	30	61.2	772	6	ABO37064	Abo37064 Human sec	778	30	61.2	772	7	ABO21810	Abo21810 Human sec
706	30	61.2	772	6	ABO41639	Abo41639 Human sec	779	30	61.2	772	7	ABO19980	Abo19980 Human sec
707	30	61.2	772	6	ABO35234	Abo35234 Human PRO	780	30	61.2	772	7	ABO24283	Abo24283 Human sec
708	30	61.2	772	6	ABM25083	Abm25083 Human sec	781	30	61.2	772	7	ABR86032	Abr86032 Human sec
709	30	61.2	772	6	ABO47475	Abo47475 Human sec	782	30	61.2	772	7	ABM10684	Abm10684 Human sec
710	30	61.2	772	6	ABO47780	Abo47780 Human sec	783	30	61.2	772	7	ABM76683	Abm76683 Human sec
711	30	61.2	772	6	ABO48390	Abo48390 Human sec	784	30	61.2	772	7	ABR89387	Abr89387 Human sec
712	30	61.2	772	6	ABO51440	Abo51440 Human PRO	785	30	61.2	772	7	ABM12514	Abm12514 Human sec
713	30	61.2	772	6	ABO51745	Abo51745 Human PRO	786	30	61.2	772	7	ABM05804	Abm05804 Human sec
714	30	61.2	772	6	ABO50525	Abo50525 Human sec	787	30	61.2	772	7	ABO34929	Abo34929 Human PRO
715	30	61.2	772	6	ABR79649	Abr79649 Human sec	788	30	61.2	772	7	ABM03005	Abm03005 Human sec
716	30	61.2	772	6	ABM16911	Abm16911 Human sec	789	30	61.2	772	7	ABM18983	Abm18983 Human sec
717	30	61.2	772	6	ABO17943	Abo17943 Human sec	790	30	61.2	772	7	ABM19288	Abm19288 Human sec
718	30	61.2	772	6	ABO20895	Abo20895 Human sec	791	30	61.2	772	7	ABO46499	Abo46499 Human PRO
719	30	61.2	772	6	ABR96854	Abr96854 Human sec	792	30	61.2	772	7	ABO49000	Abo49000 Human sec
720	30	61.2	772	6	ABM12209	Abm12209 Human sec	793	30	61.2	772	7	ABR69043	Abr69043 Human sec
721	30	61.2	772	6	ABM16301	Abm16301 Human sec	794	30	61.2	772	7	ABR89082	Abr89082 Human sec
722	30	61.2	772	6	ABM24168	Abm24168 Human sec	795	30	61.2	772	7	ABR72455	Abr72455 Human sec
723	30	61.2	772	6	ABM14649	Abm14649 Human sec	796	30	61.2	772	7	ABR74285	Abr74285 Human sec
724	30	61.2	772	6	ABM04530	Abm04530 Human sec	797	30	61.2	772	7	ABO18553	Abo18553 Human sec
725	30	61.2	772	6	ABM06719	Abm06719 Human sec	798	30	61.2	772	7	ABR80259	Abr80259 Human sec
726	30	61.2	772	6	ABM09159	Abm09159 Human sec	799	30	61.2	772	7	ABM01480	Abm01480 Human sec
727	30	61.2	772	6	ABO39199	Abo39199 Human sec	800	30	61.2	772	7	ABM02090	Abm02090 Human sec
728	30	61.2	772	6	ABM75464	Abm75464 Human sec	801	30	61.2	772	7	ABR87252	Abr87252 Human sec
729	30	61.2	772	6	ABM25388	Abm25388 Human sec	802	30	61.2	772	7	ABM12819	Abm12819 Human sec
730	30	61.2	772	6	ABM19898	Abm19898 Human sec	803	30	61.2	772	7	ABM30573	Abm30573 Human sec
731	30	61.2	772	6	ABO46804	Abo46804 Human PRO	804	30	61.2	772	7	ABM24473	Abm24473 Human sec
732	30	61.2	772	6	ABO47109	Abo47109 Human PRO	805	30	61.2	772	7	ABO29387	Abo29387 Human sec
733	30	61.2	772	6	ADA83175	Ada83175 Human sec	806	30	61.2	772	7	ABO31217	Abo31217 Human sec
734	30	61.2	772	6	ABR71540	Abr71540 Human sec	807	30	61.2	772	7	ABM14344	Abm14344 Human sec
735	30	61.2	772	6	ABR72150	Abr72150 Human sec	808	30	61.2	772	7	ABM09769	Abm09769 Human sec
736	30	61.2	772	6	ABR98489	Abr98489 Human sec	809	30	61.2	772	7	ABO38894	Abo38894 Human sec
737	30	61.2	772	6	ABO06859	Abo06859 Human sec	810	30	61.2	772	7	ABM34659	Abm34659 Human sec
738	30	61.2	772	6	ABR84812	Abr84812 Human sec	811	30	61.2	772	7	ABO51135	Abo51135 Human sec
739	30	61.2	772	6	ABR73370	Abr73370 Human sec	812	30	61.2	772	7	ABO03961	Abo03961 Human sec
741	30	61.2	772	6	ABR76464	Abr76464 Human sec	813	30	61.2	772	7	ABO10431	Abo10431 Human PRO
742	30	61.2	772	6	ABR73065	Abr73065 Human sec	814	30	61.2	772	7	ABR77674	Abr77674 Human sec
743	30	61.2	772	6	ABM18131	Abm18131 Human sec	815	30	61.2	772	7	ABR78884	Abr78884 Human sec
744	30	61.2	772	6	ABO20590	Abo20590 Human sec	816	30	61.2	772	7	ABO23978	Abo23978 Human sec
745	30	61.2	772	6	ABO25333	Abo25333 Human PRO	817	30	61.2	772	7	ABR93742	Abr93742 Human sec
746	30	61.2	772	6	ABO25638	Abo25638 Human PRO	818	30	61.2	772	7	ABM01785	Abm01785 Human sec
747	30	61.2	772	6	ABR94047	Abr94047 Human sec	819	30	61.2	772	7	ABM78208	Abm78208 Human sec
748	30	61.2	772	6	ABR79954	Abr79954 Human sec	820	30	61.2	772	7	ABR89997	Abr89997 Human sec
749	30	61.2	772	6	ABM11294	Abm11294 Human sec	821	30	61.2	772	7	ABM27523	Abm27523 Human sec
750	30	61.2	772	6	ABO32901	Abo32901 Human PRO	822	30	61.2	772	7	ABM13124	Abm13124 Human sec
751	30	61.2	772	6	ABO30607	Abo30607 Human sec	823	30	61.2	772	7	ABO31827	Abo31827 Human sec
752	30	61.2	772	6	ABO30912	Abo30912 Human sec	824	30	61.2	772	7	ABM14039	Abm14039 Human sec
753	30	61.2	772	6	ABM27218	Abm27218 Human sec	825	30	61.2	772	7	ABM08244	Abm08244 Human sec
754	30	61.2	772	6	ABM29963	Abm29963 Human sec	826	30	61.2	772	7	ABO40114	Abo40114 Human sec
	30	61.2	772	6	ABM05499	Abm05499 Human sec	827	30	61.2	772	7	ABM74549	Abm74549 Human sec

828	30	61.2	772	7	ABM33744	Abm33744 Human sec	901	30	61.2	772	8	ADF40655	Adf40655 Human sec
829	30	61.2	772	7	ABM20203	Abm20203 Human sec	902	30	61.2	772	8	ADF23599	Adf23599 Human sec
830	30	61.2	772	7	ABO48695	AbO48695 Human sec	903	30	61.2	772	8	ADF33582	Adf33582 Human sec
831	30	61.2	772	7	ABR72760	Abr72760 Human sec	904	30	61.2	772	8	ADF27049	Adf27049 Human sec
832	30	61.2	772	7	ABO15402	AbO15402 Human sec	905	30	61.2	772	8	ADF27685	Adf27685 Human sec
833	30	61.2	772	7	ABR85117	Abr85117 Human sec	906	30	61.2	772	8	ADF41279	Adf41279 Human sec
834	30	61.2	772	7	ABO15097	AbO15097 Human sec	907	30	61.2	772	8	ADF32958	Adf32958 Human sec
835	30	61.2	772	7	ABO17232	AbO17232 Human sec	908	30	61.2	772	8	ADF25324	Adf25324 Human sec
836	30	61.2	772	7	ABO19867	AbO19867 Human sec	909	30	61.2	772	8	ADF26425	Adf26425 Human sec
837	30	61.2	772	7	ABM17521	Abm17521 Human sec	910	30	61.2	772	8	ADF34214	Adf34214 Human sec
838	30	61.2	772	7	ABR85422	Abr85422 Human sec	911	30	61.2	772	8	ADF46451	Adf46451 Human sec
839	30	61.2	772	7	ABM76988	Abm76988 Human sec	912	30	61.2	772	8	ADF96068	Adf96068 Novel hum
840	30	61.2	772	7	ABO28167	AbO28167 Human sec	913	30	61.2	772	8	ADG04339	Adg04339 Novel hum
841	30	61.2	772	7	ABM22948	Abm22948 Human sec	914	30	61.2	772	8	ADG00499	Adg00499 Novel hum
842	30	61.2	772	7	ABM30268	Abm30268 Human sec	915	30	61.2	772	8	ADG82755	Adg82755 Human PRO
843	30	61.2	772	7	ABM21728	Abm21728 Human sec	916	30	61.2	772	8	ADH26036	Adh26036 Novel hum
844	30	61.2	772	7	ABM21423	Abm21423 Human sec	917	30	61.2	772	8	ADG50437	Adg50437 Human sec
845	30	61.2	772	7	ABM14954	Abm14954 Human sec	918	30	61.2	772	8	ADG63651	Adg63651 Human sec
846	30	61.2	772	7	ABO41029	AbO41029 Human sec	919	30	61.2	772	8	ADG49813	Adg49813 Human sec
847	30	61.2	772	7	ABO36759	AbO36759 Human sec	920	30	61.2	772	8	ADG51685	Adg51685 Human sec
848	30	61.2	772	7	ABO37369	AbO37369 Human sec	921	30	61.2	772	8	ADH333005	Adh333005 Human PRO
849	30	61.2	772	7	ABM75159	Abm75159 Human sec	922	30	61.2	772	8	ADG49189	Adg49189 Human sec
850	30	61.2	772	7	ABM33439	Abm33439 Human sec	923	30	61.2	772	8	ADG48565	Adg48565 Human sec
851	30	61.2	772	7	ABO46194	AbO46194 Human PRO	924	30	61.2	772	8	ADG51061	Adg51061 Human sec
852	30	61.2	772	7	ADA82541	Ada82541 Human sec	925	30	61.2	772	8	ADG59005	Adg59005 Human sec
853	30	61.2	772	7	ABM311793	Abm311793 Human sec	926	30	61.2	772	8	ADG62461	Adg62461 Human sec
854	30	61.2	772	7	ABM31183	Abm31183 Human sec	927	30	61.2	772	8	ADH25486	Adh25486 Human neu
855	30	61.2	772	7	ADB73770	Adb73770 Human PRO	928	30	61.2	772	8	ADJ54744	Adj54744 Human PRO
856	30	61.2	772	7	ADB85849	Adb85849 Human sec	929	30	61.2	772	8	ADJ64515	Adj64515 Human PRO
857	30	61.2	772	7	ABM32098	Abm32098 Human sec	930	30	61.2	772	8	ADM31411	Adm31411 Novel hum
858	30	61.2	772	7	ABM32403	Abm32403 Human sec	931	30	61.2	772	8	ADM17263	Adm17263 Human sec
859	30	61.2	772	7	ABM31488	Abm31488 Human sec	932	30	61.2	772	8	ADM36458	Adm36458 Novel hum
860	30	61.2	772	7	ABM30878	Abm30878 Human sec	933	30	61.2	772	8	ADM40263	Adm40263 Novel hum
861	30	61.2	772	7	ADB76486	Adb76486 Human PRO	934	30	61.2	772	8	ADL07097	Adl07097 Human sec
862	30	61.2	772	7	ADC43912	Adc43912 Human sec	935	30	61.2	772	8	ADN37871	Adn37871 Novel hum
863	30	61.2	772	7	ADC61672	Adc61672 Human sec	936	30	61.2	772	8	ADP55976	Adp55976 Human PRO
864	30	61.2	772	7	ADC63636	Adc63636 Human sec	937	30	61.2	772	8	ADU50219	Adu50219 PRO941, S
865	30	61.2	772	7	ADC66736	Adc66736 Human sec	938	30	61.2	772	9	ADW49498	Adw49498 PRO941 pr
866	30	61.2	772	7	ADC68860	Adc68860 Human sec	939	30	61.2	772	9	ADZ52159	Adz52159 Human sec
867	30	61.2	772	7	ADC62920	Adc62920 Human sec	940	30	61.2	774	8	ADS28353	Ads28353 Bacterial
868	30	61.2	772	7	ADC67985	Adc67985 Human sec	941	30	61.2	789	6	ADS44794	Ads44794 Bacterial
869	30	61.2	772	7	ADC41305	Adc41305 Human sec	942	30	61.2	791	6	ABU25481	Abu25481 Protein e
870	30	61.2	772	7	ADC67360	Adc67360 Human sec	943	30	61.2	791	6	ABU48892	Abu48892 Protein e
871	30	61.2	772	7	ADC62296	Adc62296 Human sec	944	30	61.2	798	6	ABU38404	Abu38404 Protein e
872	30	61.2	772	7	ADC41929	Adc41929 Human sec	945	30	61.2	798	6	ABU40004	Abu40004 Protein e
873	30	61.2	772	7	ADD05579	Add05579 Human sec	946	30	61.2	798	6	ADS24706	Ads24706 Bacterial
874	30	61.2	772	7	ADE49298	Ade49298 Human sec	947	30	61.2	799	8	ADS27901	Ads27901 Bacterial
875	30	61.2	772	7	ADE35352	Ade35352 Human sec	948	30	61.2	805	6	ABU22655	Abu22655 Protein e
876	30	61.2	772	7	ADE16466	Ade16466 Human sec	949	30	61.2	807	6	ABU21205	Abu21205 Protein e
877	30	61.2	772	7	ADD73081	Add73081 Human sec	950	30	61.2	807	6	ABU19894	Abu19894 Protein e
878	30	61.2	772	7	ADD72439	Add72439 Human sec	951	30	61.2	807	8	ADN74391	Adn74391 Thale cre
879	30	61.2	772	7	ADE17090	Ade17090 Human sec	952	30	61.2	808	8	ADN26459	Adn26459 Bacterial
880	30	61.2	772	7	ADF47104	Adf47104 Human sec	953	30	61.2	811	3	AAG35867	Aag35867 Arabidops
881	30	61.2	772	7	ADG02574	Adg02574 Novel hum	954	30	61.2	817	6	ABU22957	Abu22957 Protein e
882	30	61.2	772	7	ADG01281	Adg01281 Novel hum	955	30	61.2	820	6	ABP79190	Abp79190 N. gonorr
883	30	61.2	772	7	ADF95456	Adf95456 Novel hum	956	30	61.2	820	6	ABU37137	Abu37137 Protein e
884	30	61.2	772	7	ADG12271	Adg12271 Novel hum	957	30	61.2	820	6	ABU37980	Abu37980 Protein e
885	30	61.2	772	7	ADH08931	Adh08931 Human PRO	958	30	61.2	820	6	ABO81458	AbO81458 Pseudomon
886	30	61.2	772	7	ADG52861	Adg52861 Human sec	959	30	61.2	820	8	ADP08282	Adp08282 Neisseria
887	30	61.2	772	7	ADG60181	Adg60181 Human sec	960	30	61.2	824	7	ADB64361	Adb64361 Human pro
888	30	61.2	772	7	ADG63802	Adg63802 Human sec	961	30	61.2	826	6	ABU28255	Abu28255 Protein e
889	30	61.2	772	7	ADL32712	Adl32712 Novel hum	962	30	61.2	828	5	ABP69286	Abp69286 Human cel
890	30	61.2	772	7	ADM30246	Adm30246 Novel hum	963	30	61.2	828	6	ABP58225	Abp58225 Human cel
891	30	61.2	772	7	ADI60941	Adi60941 Human sec	964	30	61.2	842	8	ADQ65365	Adq65365 Novel hum
892	30	61.2	772	8	ADE48598	Ade48598 Human sec	965	30	61.2	848	5	ABG79689	Abg79689 Tumour in
893	30	61.2	772	8	ADE74243	Ade74243 Human sec	966	30	61.2	855	7	ABM90485	Abm90485 Rice abio
894	30	61.2	772	8	ADE41216	Ade41216 Human sec	967	30	61.2	859	4	ABG21946	Abg21946 Novel hum
895	30	61.2	772	8	ADE74855	Ade74855 Human sec	968	30	61.2	881	6	ABU48605	Abu48605 Protein e
896	30	61.2	772	8	ADE89699	Ade89699 Human sec	969	30	61.2	889	8	ADN23791	Adn23791 Bacterial
897	30	61.2	772	8	ADF61339	Adf61339 Human sec	970	30	61.2	900	8	ADP26907	Adp26907 Human N-c
898	30	61.2	772	8	ADF40031	Adf40031 Human sec	971	30	61.2	905	5	ABG93053	Abg93053 S. cerevi
899	30	61.2	772	8	ADF45827	Adf45827 Human sec	972	30	61.2	905	6	ABR53130	Abr53130 Protein s
900	30	61.2	772	8	ADF24223	Adf24223 Human sec	973	30	61.2	905	7	ADK62564	Adk62564 Disease t

974 30 61.2 905 8 ADQ20653 Adq20653 Human sof
975 30 61.2 906 3 AAY70741 Aay70741 Human N-c
976 30 61.2 906 4 ABG14316 Abg14316 Novel hum
977 30 61.2 906 5 ABB81474 Abb81474 Human N-c
978 30 61.2 906 5 ABB57233 Abb57233 Mouse isc
979 30 61.2 906 6 ABR58643 Abr58643 Human can
980 30 61.2 906 6 ABR47406 Abr47406 Breast ca
981 30 61.2 906 7 ADE55478 Ade55478 Human Pro
982 30 61.2 906 7 ADE55482 Ade55482 Human Pro
983 30 61.2 906 7 ADE55488 Ade55488 Rat Prote
984 30 61.2 906 7 ADE55476 Ade55476 Rat Prote
985 30 61.2 906 7 ADE55486 Ade55486 Human Pro
986 30 61.2 906 7 ADE55484 Ade55484 Rat Prote
987 30 61.2 906 7 ADE55490 Ade55490 Human Pro
988 30 61.2 906 7 ADE55480 Ade55480 Rat Prote
989 30 61.2 906 7 ADN95503 Adn95503 Human BEC
990 30 61.2 906 8 ADR69159 Adr69159 Human N-c
991 30 61.2 906 8 ADN23873 Adn23873 Bacterial
992 30 61.2 906 9 ADY70663 Ady70663 Human nic
993 30 61.2 906 9 ADZ26571 Adz26571 Human N-c
994 30 61.2 912 5 ABB81471 Abb81471 Chicken N
995 30 61.2 913 8 ADU73677 Adu73677 Mouse cad
996 30 61.2 916 7 ADN38968 Adn38968 Cancer/an
997 30 61.2 916 8 ADU73673 Adu73673 Human cad
998 30 61.2 922 3 AAG35866 Aag35866 Arabidops
999 30 61.2 947 4 ABB59355 Abb59355 Drosophil
1000 30 61.2 976 5 ABP35659 Abp35659 Fungal ZB

ALIGNMENTS

RESULT 1
AAY22358
ID AAY22358 standard; peptide; 10 AA.
XX
AC AAY22358;
XX
DT 27-SEP-1999 (first entry)
DE
DE TPO receptor binding peptide sequence, SEQ ID NO. 9.
XX
XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
XX US5932546-A.
XX
XX 03-AUG-1999.
PD
XX 04-OCT-1996; 96US-00726464.
PF
XX 04-OCT-1996; 96US-00726464.
PR
XX (GLAX) GLAXO WELLCOME INC.
PA
XX
XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
XX WPI; 1999-457122/38.
DR
XX New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
PT
XX
XX Disclosure; Col 13-14; 36pp; English.
PS
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db |||||
1 ESVREQVMKY 10

RESULT 2
AAB16982
ID AAB16982 standard; peptide; 10 AA.
XX
AC AAB16982;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:38.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
XX
XX 04-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US025044.
PF
XX 23-OCT-1998; 98US-0105371P.
PR
XX 22-OCT-1999; 99US-00428082.
XX
XX (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheetham J, Boone TC;
PI
XX WPI; 2000-350702/30.
DR
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 19; Page 208; 608pp; English.
PS
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | | | |
Db 1 ESVREQVMKY 10

RESULT 3
ABB72868
ID ABB72868 standard; peptide; 10 AA.
XX
AC ABB72868;
XX
DT 05-APR-2002 (first entry)
DE TPO mimetic peptide SEQ ID NO:38.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | | | |
Db 1 ESVREQVMKY 10

RESULT 4
ADJ73019
ID ADJ73019 standard; peptide; 10 AA.
XX
AC ADJ73019;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 473.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX WPI; 2003-804237/75.
DR
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 473; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 49; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | | | |
Db 1 ESVREQVMKY 10

RESULT 5
ADJ52654
ID ADJ52654 standard; peptide; 10 AA.
XX
AC ADJ52654;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID473.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX
OS Unidentified.
OS Synthetic.
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX
PS Claim 2; SEQ ID NO 473; 129pp; English.

XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 49; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | | | |
Db 1 ESVREQVMKY 10

RESULT 6
ADJ51615
ID ADJ51615 standard; peptide; 10 AA.
XX
AC ADJ51615;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID473.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.

XX
PS Claim 14; SEQ ID NO 473; 123pp; English.

XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | | | |
Db 1 ESVREQVMKY 10

RESULT 7
ABU36234
ID ABU36234 standard; protein; 291 AA.

AC ABU36234;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #21761.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

KW Mycoplasma pneumoniae.

OS WO200277183-A2.

PN 03-OCT-2002.

PD 21-MAR-2002; 2002WO-US009107.

PF 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

PI Walli D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA40104.

XX New antisense nucleic acids, useful for identifying proteins or screening
CC for homologous nucleic acids required for cellular proliferation to
CC isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 64158; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 291 AA;

Query Match 77.6%; Score 38; DB 6; Length 291;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : | | | : | |
Db 191 EALREQIISKY 200

RESULT 8
ADN99878
ID ADN99878 standard; protein; 216 AA.

XX ADN99878;

DT 29-JUL-2004 (first entry)

DE Novel human protein sequence #694.

XX anti-inflammatory; dermatological; neuroprotective; immunomodulator;
KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;
KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
KW early aging; hormonal imbalance; ischemic heart disease;
KW ulcerative colitis.

XX Homo sapiens.

OS WO2004038003-A2.

PN 06-MAY-2004.

PF 24-OCT-2003; 2003WO-US033947.

XX 25-OCT-2002; 2002US-0421061P.

PR 25-OCT-2002; 2002US-0421080P.

PR 25-OCT-2002; 2002US-0421552P.

PR 25-OCT-2002; 2002US-0421614P.

PR 30-OCT-2002; 2002US-0422177P.

PR 30-OCT-2002; 2002US-0422178P.

PR 15-NOV-2002; 2002US-0426355P.

PR 15-NOV-2002; 2002US-0426384P.

PR 15-NOV-2002; 2002US-0426394P.

PR 15-NOV-2002; 2002US-0426430P.

PR 15-NOV-2002; 2002US-0426916P.

PR 27-NOV-2002; 2002US-0429224P.

PR 27-NOV-2002; 2002US-0429275P.

PR 27-NOV-2002; 2002US-0429302P.

PR 27-NOV-2002; 2002US-0429326P.

PR 27-NOV-2002; 2002US-0429651P.

PR 04-DEC-2002; 2002US-0430645P.

PR 04-DEC-2002; 2002US-0430651P.

PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
DR WPI; 2004-365511/34.
DR N-PSDB; ADN99094.
XX
PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
PS Claim 14; SEQ ID NO 1478; 532pp; English.
XX
CC The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
SQ Sequence 216 AA;

Query Match 75.5%; Score 37; DB 8; Length 216;

Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db |::|||:
72 EKIREQIKY 81

RESULT 9
AAY20122
ID AAY20122 standard; protein; 384 AA.
XX
AC AAY20122;
XX
DT 19-JUL-1999 (first entry)
XX
DE B. burgdorferi antigenic protein, f6-27.nt.
XX
KW Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS Borrelia burgdorferi.
XX
PN WO9859071-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US012718.
XX
PR 20-JUN-1997; 97US-0050359P.
PR 22-JUL-1997; 97US-0053344P.
PR 22-JUL-1997; 97US-0053377P.
PR 03-SEP-1997; 97US-0057483P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMMUNE INC.
XX
PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX
DR WPI; 1999-189980/16.
DR N-PSDB; AAX61819.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases caused
PT by Borrelia, particularly Lyme disease.
XX
PS Claim 12; Page 206; 275pp; English.
XX
CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus
XX
SQ Sequence 384 AA;

Query Match 75.5%; Score 37; DB 2; Length 384;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db |::|||:
33 ESLKEQMKY 42

RESULT 10
AAY06993
ID AAY06993 standard; protein; 429 AA.
XX
AC AAY06993;
XX
DT 02-JUL-1999 (first entry)

XX DE Rat eukaryotic translation initiation factor (eIF-5).

XX KW Cancer associated antigen; diagnosis; research; treatment; human;

XX KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;

XX KW prostate cancer.

XX OS Rattus sp.

XX PN WO9904265-A2.

XX PD 28-JAN-1999.

XX PF 15-JUL-1998; 98WO-US014679.

XX PR 17-JUL-1997; 97US-00896164.

XX PR 10-OCT-1997; 97US-0061599P.

XX PR 10-OCT-1997; 97US-0061765P.

XX PR 10-OCT-1997; 97US-00948705.

XX PR 11-OCT-1997; 97GB-00021697.

XX PR 22-JUN-1998; 98US-00102322.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;

XX PI O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;

XX DR WPI; 1999-132448/11.

XX DR N-PSDB; AAX40194.

XX XX New isolated cancer associated nucleic acids and polypeptides - isolated using sera from cancer patients, used to develop products for the diagnosis, monitoring or treatment of cancers.

XX PS Example 8; Page 771-772; 787pp; English.

XX CC The invention relates to a method for diagnosing a disorder characterised by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a biological sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the interaction between the agent and the NAM or the expression product as a determination of the disorder. The products and methods can be used in the diagnosis, monitoring, research, or treatment of conditions characterised by the expression of various cancer associated antigens. The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and lung cancer

XX SQ Sequence 429 AA;

Query Match 75.5%; Score 37; DB 2; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 285 EKIREQIKKY 294

RESULT 11
ADE60677
ID ADE60677 standard; protein; 429 AA.
XX ADE60677;
AC
XX 29-JAN-2004 (first entry)
DT
DE Rat Protein Q07205, SEQ ID NO 6589.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX XX WPI; 2003-268312/26.

XX DR GENBANK; Q07205.

XX PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 429 AA;

Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 285 EKIREQIKKY 294

RESULT 12
ADD45400
ID ADD45400 standard; protein; 429 AA.
XX
AC ADD45400;
XX
DT 29-JAN-2004 (first entry)

XX DE Rat Protein Q07205, SEQ ID NO 10833.
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q07205.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 429 AA;
Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
RESULT 13
ADE60681
ID ADE60681 standard; protein; 429 AA.

XX AC ADE60681;
XX DT 29-JAN-2004 (first entry)
XX DE Rat Protein Q07205, SEQ ID NO 6593.
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q07205.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 429 AA;
Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294

RESULT 14
ADE61162
ID ADE61162 standard; protein; 429 AA.
XX
AC ADE61162;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein Q07205, SEQ ID NO 7080.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
XX WO2003016475-A2.
PN
PD 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; Q07205.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 429 AA;
Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| :|||: ||
Db 285 EKIREQIKKY 294
RESULT 15
ADE61164
ID ADE61164 standard; protein; 431 AA.
XX
AC ADE61164;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P55010, SEQ ID NO 7082.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| :|||: ||
Db 287 EKIREQIKKY 296

RESULT 16
ADE60679
ID ADE60679 standard; protein; 431 AA.
XX
AC ADE60679;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P55010, SEQ ID NO 6591.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| :|||: ||
Db 287 EKIREQIKKY 296

RESULT 17
ADD45402
ID ADD45402 standard; protein; 431 AA.
XX
AC ADD45402;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P55010, SEQ ID NO 10835.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 18
ADE60683
ID ADE60683 standard; protein; 431 AA.
XX
AC ADE60683;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P55010, SEQ ID NO 6595.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 19
ADN04323
ID ADN04323 standard; protein; 431 AA.
XX
AC ADN04323;

XX 01-JUL-2004 (first entry)
DT
XX Antipsoriatic protein sequence #356.
DE
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
KW
XX Homo sapiens.

OS
XX WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
PI
XX WPI; 2004-305105/28.
DR N-PSDB; ADN04322.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 9; SEQ ID NO 717; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 20
ABM81587
ID ABM81587 standard; protein; 431 AA.

XX AC ABM81587;

XX DT 18-NOV-2004 (first entry)

XX DE Tumour-associated antigenic target (TAT) polypeptide PRO82391, SEQ:4099.

KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX DR WPI; 2004-347921/32.

XX DR N-PSDB; ACN39755.

XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

PS Claim 12; SEQ ID NO 4099; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention

XX SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 21

ADP54723

ID ADP54723 standard; protein; 431 AA.

XX AC ADP54723;

XX DT 18-NOV-2004 (first entry)

XX DE Human PRO protein sequence SEQ ID NO:699.

XX human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.

XX OS Homo sapiens.

XX PN WO2004039956-A2.

XX PD 13-MAY-2004.

XX PF 28-OCT-2003; 2003WO-US034381.

XX PR 29-OCT-2002; 2002US-0422472P.

XX PA (GETH) GENENTECH INC.

XX PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

PI Wood WI, Wu TD;

XX WPI; 2004-376182/35.

XX N-PSDB; ADP54722.

PT New PRO polynucleotides and polypeptides, useful in diagnosing
PT and treating an immune related disease, e.g. systemic lupus
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT stimulating an immune response.

XX Claim 1; SEQ ID NO 699; 3009pp; English.

PS The present invention describes an isolated PRO nucleic acid (I). Also
XX described: (1) a vector comprising (I); (2) a host cell comprising the
CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC antibody which specifically binds to a polypeptide of (4); (7) a
CC composition of matter comprising a polypeptide of (4), an agonist or
CC antagonist of the polypeptide or an antibody that binds to the
CC polypeptide in combination with a carrier; (8) an article of manufacture
CC comprising a container, a label on the container and a composition of
CC matter of (7); (9) a method of treating an immune related disease in a
CC mammal; (10) a method for determining the presence of a PRO polypeptide
CC in a sample suspected of having the polypeptide; (11) a method of
CC diagnosing an immune related disease or an inflammatory immune response
CC in mammal; (12) a method of identifying a compound that inhibits or
CC mimics the activity of or expression of a gene encoding a PRO polypeptide
CC ; and (13) a method of stimulating the immune response in a mammal. The
CC PRO sequences have antiallergic, antianaemic, antiarthritic,
CC antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC virucide activities, and can be used in gene therapy. The nucleic acid
CC (I) and the encoded polypeptides, compositions, kits and methods are

CC useful in diagnosing and treating an immune related disease and in
CC stimulating an immune response. The present sequence represents a human
CC PRO protein from the present invention.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 22
ADX05561
ID ADX05561 standard; protein; 431 AA.
XX
AC ADX05561;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 126.
XX
KW cytotostatic; cyclin-dependent kinase; cdk; biomarker.
XX
OS Homo sapiens.
XX
PN WO2005012875-A2.
XX
PD 10-FEB-2005.
XX
PF 29-JUL-2004; 2004WO-US024424.
XX
PR 29-JUL-2003; 2003US-0490890P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX
XX WPI; 2005-163068/17.
DR N-PSDB; ADX05560.
XX
XX Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 126; 141pp; English.
XX
XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence represents a biomarker used in the method of the invention.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 9; Length 431;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 23
ADY15998
ID ADY15998 standard; protein; 431 AA.
XX
AC ADY15998;
XX
DT 05-MAY-2005 (first entry)
XX
DE PRO polypeptide SEQ ID NO 1804.
XX
KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2005016962-A2.
XX
PD 24-FEB-2005.
XX
PF 11-AUG-2004; 2004WO-US026249.
XX
PR 11-AUG-2003; 2003US-0493546P.
XX
PA (GETH) GENENTECH INC.
XX
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX
DR WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS Claim 8; SEQ ID NO 1804; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.
XX
SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 9; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 24
AAB43508
ID AAB43508 standard; protein; 433 AA.
XX
AC AAB43508;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:953.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC77717.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 1519-1521; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 433 AA;

Query Match 75.5%; Score 37; DB 3; Length 433;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 289 EKIREQIKKY 298

RESULT 25
ADT56493
ID ADT56493 standard; protein; 514 AA.
XX
AC ADT56493;
XX

DT 13-JAN-2005 (first entry)
XX Plant polypeptide, SEQ ID 6570.
DE
XX
KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
DR WPI; 2004-757369/74.
XX
PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
PS Claim 2; SEQ ID NO 6570; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 514 AA;

Query Match 75.5%; Score 37; DB 8; Length 514;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 436 EEIREQALKY 445

RESULT 26
AAU30330
ID AAU30330 standard; protein; 590 AA.

XX AC AAU30330;
XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #821.

XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.

XX PS Claim 20; Page 278; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

XX SQ Sequence 590 AA;

Query Match 75.5%; Score 37; DB 4; Length 590;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 435 EKIREQIKKY 444

RESULT 27
AAY22378
ID AAY22378 standard; peptide; 15 AA.

XX AC AAY22378;

XX DT 27-SEP-1999 (first entry)

XX DE TPO receptor binding peptide sequence, SEQ ID NO. 29.

XX KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX OS Synthetic.

XX PN US5932546-A.

XX PD 03-AUG-1999.

XX PF 04-OCT-1996; 96US-00726464.

XX PR 04-OCT-1996; 96US-00726464.

XX PA (GLAX) GLAXO WELLCOME INC.

XX PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

XX PI Cwirla SE, Johnson SS;

XX DR WPI; 1999-457122/38.

XX PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.

XX PS Disclosure; Col 13-14; 36pp; English.

XX CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration

XX SQ Sequence 15 AA;

Query Match 73.5%; Score 36; DB 2; Length 15;
Best Local Similarity 70.0%; Pred. No. 5.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 4 ESVRDQVMKF 13

RESULT 28
ADC89126
ID ADC89126 standard; protein; 150 AA.

XX AC ADC89126;

XX DT 01-JAN-2004 (first entry)

XX DE Ribosomal protein similar to FCWP1 #1342.

XX KW Antifungal protein; ribosomal protein; FCWP1; AlyAFP;
KW plant fungal infection; Alternaria; Ascochyta; Botrytis; Cercospora;
KW Colletotrichum; Diplodia; Fusarium; Gaemanomyces; Helminthosporium;

KW Macrophomina; Mycosphaerella; Nectria; Peronospora; Phoma;
KW Phymatotrichum; Phytophthora; Plasmopara; Podosphaera; Puccinia; Puthium;
KW Pyrenophora; Pycularia; Pythium; Rhizoctonia; Sclerotium; Sclerotinia;
KW Septoria; Thielaviopsis; Venturia; Verticillium.
XX Unidentified.
OS
XX US6573361-B1.
PN
XX
PD 03-JUN-2003.
XX
PF 07-DEC-2000; 2000US-00732210.
XX
PR 07-DEC-1999; 99US-0169340P.
XX
PR 07-DEC-1999; 99US-0169513P.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Bunkers GJ, Liang J, Mittanck CA, Seale JW, Wu YS;
XX WPI; 2003-754558/71.
DR
XX Novel antifungal protein FCWP1, isolated from Fusarium culmorum, useful
PT for controlling fungal infections in plants.
PT
XX
PS Example 21; SEQ ID NO 1379; 27pp; English.
XX
CC The invention relates to an isolated antifungal ribosomal protein from
CC fusarium culmorum, FCWP1. Also included is a fusion protein between the
CC signal peptide of the antifungal protein AlyAPP from Alyssum and FCWP1,
CC encoded by the nucleic acid appearing as ADC87758. The FCWP1 proteins are
CC useful for controlling fungal infections in plants, such as those caused
CC by Alternaria (e.g. Alternaria brassicola , Alternaria solani),
CC Ascochyta (e.g. Ascochyta pisi); Botrytis (e.g. Botrytis cinerea),
CC Cercospora (e.g. Cercospora kikuchii , Cercospora zaea-maydis),
CC Colletotrichum (e.g. Colletotrichum lindemuthianum), Diplodia (e.g.
CC Diplodia maydis), Fusarium (e.g. Fusarium nivale , Fusarium oxysporum ,
CC Fusarium graminearum , Fusarium culmorum , Fusarium solani , Fusarium
CC moniliforme , Fusarium roseum), Gaeumanomyces (e.g. Gaeumanomyces
CC graminis f.sp. tritici), Helminthosporium (e.g. Helminthosporium turcicum
CC , Helminthosporium carbonum , Helminthosporium maydis), Macrospora
CC (e.g. Macrospora phaseolina , Maganaporthe grisea), Mycosphaerella
CC (e.g. Mycosphaerella figlensis), Nectria (Nectria haematococca), Phoma
CC Peronospora (e.g. Peronospora manshurica , Peronospora tabacina), Phoma
CC (e.g. Phoma betae), Phymatotrichum (e.g. Phymatotrichum omnivorum),
CC Phytophthora (e.g. Phytophthora cinnamomi , Phytophthora cactorum ,
CC Phytophthora phaseoli , Phytophthora parasitica , Phytophthora
CC citrophthora , Phytophthora megasperma f.sp. sojae, Phytophthora
CC infestans), Plasmopara (e.g. Plasmopara viticola), Podosphaera (e.g.
CC Podosphaera leucotricha), Puccinia (e.g. Puccinia sorghi , Puccinia
CC striiformis , Puccinia graminis f.sp. tritici, Puccinia asparagi ,
CC Puccinia recondita , Puccinia arachidis), Puthium (e.g. Puthium
CC aphanidermatum), Pyrenophora (e.g. Pyrenophora tritici-repentens),
CC Pycularia (e.g. Pycularia oryzae), Pythium (e.g. Pythium ultimum),
CC Rhizoctonia (e.g. Rhizoctonia solani , Rhizoctonia cerealis), Sclerotium
CC (e.g. Sclerotium rolfsii), Sclerotinia (e.g. Sclerotinia sclerotiorum),
CC Septoria (e.g. Septoria lycopersici , Septoria glycines , Stagonospora
CC nodorum / Phaeosphaeria nodorum , Septoria tritici), Thielaviopsis (e.g.
CC Thielaviopsis basicola), Uncinula (e.g. Uncinula necator), Venturia
CC (e.g. Venturia inaequalis) or Verticillium (e.g. Verticillium dahliae ,
CC Verticillium albo-atrum). Mutations in the proteolytic consensus
CC sequences contained within FCWP1 provides improved stability of its
CC antifungal activity. Also disclosed are ribosomal proteins with similar
CC PI (>7) and molecular weight (<20kDa) to FCWP1, which may act as
CC antifungal proteins. The present sequence represents one of the ribosomal
CC proteins similar to FCWP1. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=6573361B1.

XX Sequence 150 AA;

Query Match 73.5%; Score 36; DB 7; Length 150;

Best Local Similarity 70.0%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
||| ||: ||
Db 30 ESVIEQIVKY 39
RESULT 29
ABG93066
ID ABG93066 standard; protein; 151 AA.
XX
AC ABG93066;
XX
DT 21-NOV-2002 (first entry)
XX
DE S. cerevisiae BAX-associated protein fragment SEQ ID 90.
XX
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KW neurodegeneration; cell death.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200264766-A2.
XX
PD 22-AUG-2002.
XX
PF 21-DEC-2001; 2001WO-EP015398.
XX
PR 22-DEC-2000; 2000EP-00870318.
PR 04-JAN-2001; 2001EP-00870002.
PR 09-JAN-2001; 2001EP-00870003.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
DR WPI; 2002-667002/71.
XX N-PSDB; ABQ76332.
XX
PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
PS Claim 36; Fig 1; 344pp; English.
XX
CC This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide; immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides , or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polypeptide associated with the Bax gene
CC described in the disclosure of the invention
XX
SQ Sequence 151 AA;

Query Match 73.5%; Score 36; DB 5; Length 151;

Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| ||:|
Db 31 ESVIEQIVKY 40

RESULT 30
ADK64530
ID ADK64530 standard; protein; 151 AA.

XX AC ADK64530;

XX XX 06-MAY-2004 (first entry)

XX DE Disease treating protein complex-derived protein #1110.

XX KW protein complex; drug target; diagnosis.

XX XX Unidentified.

XX PN EP1338608-A2.

XX PD 27-AUG-2003.

XX PF 20-DEC-2002; 2002EP-00102902.

XX PR 20-DEC-2001; 2001EP-00130253.

XX PA (CELL-) CELLZOME AG.

XX PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
PI Michon A, Leutwein C, Rick J;

XX DR WPI; 2003-638460/61.
DR N-PSDB; ADK64531.

XX PT New proteins and protein complexes from eukaryotes, useful as targets in
PT drug screening, or in diagnosing or screening for the presence of a
PT disease or disorder, or a predisposition for developing a disease or
PT disorder in a subject.

XX PS Disclosure; SEQ ID NO 2219; 13pp; English.

XX CC The invention relates to novel protein complexes comprising a first and a
XX second protein, or its derivative, fragment, homologue or variant. The
XX proteins are selected from given protein complexes, which are not defined
XX in the specification. The variants are encoded by nucleic acids that
XX hybridize to the nucleic acids encoding the proteins under low stringency
XX conditions. The protein complexes are useful as targets for an active
XX agent of a pharmaceutical. These protein complexes are particularly
XX useful as drugs targets for the treatment or preventing of a disease or
XX disorder. The complexes and methods above are useful in diagnosing or
XX screening for the presence of a disease or disorder in a subject. These are also useful
XX for developing a disease or disorder in a subject. These are also useful
XX in screening for a drug for treatment or prevention of a disease or
XX disorder. The molecule that modulates the amount, activity or protein
XX components of the complex is useful for the manufacture of a medicament
XX for the treatment or prevention of a disease or disorder. This sequence
XX corresponds to a protein of the invention. (Note: the sequence data for
XX this patent did not form part of the printed specification but was
XX obtained from the EPO in electronic format).

XX SQ Sequence 151 AA;

Query Match 73.5%; Score 36; DB 7; Length 151;
Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| ||:|

Db 31 ESVIEQIVKY 40

RESULT 31
AAU67591
ID AAU67591 standard; protein; 339 AA.

XX AC AAU67591;

XX XX 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #28487.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

XX PR 02-JUN-2000; 2000US-0208841P.

XX PR 07-JUL-2000; 2000US-0216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59583.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX PS Example 1; SEQ ID NO 28786; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 339 AA;

Query Match 73.5%; Score 36; DB 4; Length 339;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|

Db 104 ESIKDQVSKY 113

RESULT 32

ABM64110

ID ABM64110 standard; protein; 339 AA.

XX

AC ABM64110;

XX

DT 20-OCT-2003 (first entry)

XX

DE Propionibacterium acnes predicted ORF-encoded polypeptide #28786.

XX

KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.

XX

OS Propionibacterium acnes.

XX

PN WO2003033515-A1.

XX

PD 24-APR-2003.

XX

PF 11-OCT-2002; 2002WO-US032727.

XX

PR 15-OCT-2001; 2001US-00978825.

XX

PA (CORI-) CORIXA CORP.

XX

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX

DR WPI; 2003-381789/36.

DR N-PSDB; ACF64512.

XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX

PS Example 1; SEQ ID NO 28786; 1481pp; English.

XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or

CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a

CC patient; and a method for inhibiting the development of P. acnes in a

CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes

CC protein. The polynucleotides can also be used as probes or primers for

CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a polypeptide predicted to be encoded by an ORF (open

CC reading frame) contained within the P. acnes polynucleotides of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 339 AA;

Query Match 73.5%; Score 36; DB 6; Length 339;

Best Local Similarity 60.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ESVREQVMKY 10

Db 104 ESIKDQVSKY 113

RESULT 33

ABM65793

ID ABM65793 standard; protein; 339 AA.

XX

AC ABM65793;

XX

DT 20-OCT-2003 (first entry)

XX

DE Propionibacterium acnes immunogenic polypeptide #30469.

XX

KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine; immunogenic.

XX

OS Propionibacterium acnes.

XX

PN WO2003033515-A1.

XX

PD 24-APR-2003.

XX

PF 11-OCT-2002; 2002WO-US032727.

XX

PR 15-OCT-2001; 2001US-00978825.

XX

PA (CORI-) CORIXA CORP.

XX

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX

DR WPI; 2003-381789/36.

XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX

PS Claim 7; SEQ ID NO 30469; 1481pp; English.

XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or

CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a

CC patient; and a method for inhibiting the development of P. acnes in a

CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes

CC protein. The polynucleotides can also be used as probes or primers for

CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a specifically claimed P. acnes polypeptide which is

CC thought to contain an immunogenic region. Note: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 339 AA;

Query Match 73.5%; Score 36; DB 6; Length 339;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||::|||
Db 104 ESIKQVSKY 113

RESULT 34
ABG28636
ID ABG28636 standard; protein; 914 AA.

XX AC ABG28636;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #28627.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS92823.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 58995; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological actions. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 914 AA;

SQ Query Match

73.5%; Score 36; DB 4; Length 914;

Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
||::|||
Db 334 ESVKEQVMR 342

RESULT 35
ADQ97432
ID ADQ97432 standard; protein; 1250 AA.

XX AC ADQ97432;

DT 07-OCT-2004 (first entry)

DE Mouse cancer associated sequence MP08-047, SEQ ID 409.

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.

OS Mus musculus.

PN WO2004060304-A2.

XX 22-JUL-2004.

XX 22-DEC-2003; 2003WO-US041389.

XX 27-DEC-2002; 2002US-00330773.

PA (SAGR-) SAGRES DISCOVERY INC.

PI Morris DW, Malandro MS;

DR WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.

PS Claim 1; SEQ ID NO 409; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1250 AA;

Query Match 73.5%; Score 36; DB 8; Length 1250;
Best Local Similarity 80.0%; Pred. No. 6.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||
Db 589 ESVRAQVEKY 598

RESULT 36
ADQ90750
ID ADQ90750 standard; protein; 2110 AA.

XX AC ADQ90750;

DT 21-OCT-2004 (first entry)

XX Mouse acetyl-coenzyme A carboxylase - ACCL.

XX mouse; carboxyltransferase domain; CT domain;
KW acetyl-coenzyme A carboxylase; acetyl-CoA carboxylase;
KW metabolic syndrome; diabetes; obesity; cardiovascular disease;

KW	atherosclerosis; depression; cancer; hyperlipidaemia; dyslipidaemia;	PA	(BIOL-) FORSCHUNGSINSTITUT BIOLOGIE LANDWIRTSCHA.
KW	hypertension; hyperuricaemia; renal disfunction;	XX	
KW	crystallisable composition; enzyme; murine; ACC1.	PI	Seyfert HM;
XX		XX	
OS	Mus musculus.	DR	WPI; 2001-258968/27.
XX		DR	N-PSDB; AAF88002.
PN	WO2004063715-A2.	XX	
XX		PT	New nucleic acid encoding bovine acetyl coenzyme A carboxylase alpha and
PD	29-JUL-2004.	PT	its promoter, for milk-specific production of proteins and for regulating
XX		PT	fat content of milk.
PF	09-JAN-2004; 2004WO-US0000585.	XX	
XX		PS	Claim 2c; Page 24-30; 44pp; German.
PR	10-JAN-2003; 2003US-0439383P.	XX	
PR	31-MAR-2003; 2003US-0459464P.	CC	This invention describes a novel milk gland-specific promoter of the
PR	31-JUL-2003; 2003US-0491640P.	CC	bovine acetyl coenzyme A carboxylase alpha gene (Accalpha), or its
PR	27-OCT-2003; 2003US-0514636P.	CC	fragments, which are used to control expression of foreign genes. When
PR	09-JAN-2004; 2004US-00514636.	CC	the promoter (or the Accalpha structural gene) is replaced, at least in
XX		CC	part, by a sequence that is altered by deletion or substitution, then
XX		CC	expression of Accalpha in the milk gland is reduced and the milk produced
PA	(UYCO) UNIV COLUMBIA NEW YORK.	CC	(by cattle, sheep or goats) has reduced fat content. Analysis of the
XX		CC	polymorphic 933-966 region of (Accalpha) is useful for genotyping
PI	Tong L, Zhang H, Yang Z;	CC	animals, producing a genotype that is directly correlated with Accalpha
XX		CC	expression during lactation and with fat content of the milk. This makes
DR	WPI; 2004-571486/55.	CC	it possible to generate populations of cattle that produce milk of high
XX		CC	or low fat content by classical breeding methods. Accalpha is a lactation
PT	New crystallizable composition comprising a carboxyltransferase domain of	CC	-specific, inducible promoter for expressing foreign proteins in the milk
PT	acetyl-coenzyme A (CoA) carboxylase, for treating e.g. diabetes, obesity,	CC	and, when modified, results in milk of reduced fat content, which
PT	cardiovascular disease, atherosclerosis, or cancer.	CC	facilitates recovery of proteins. This sequence represents a fragment
XX		CC	from bovine Accalpha which contains a fragment of exon 5A exon 6 and exon
PS	Example 4; SEQ ID NO 13; 195pp; English.	CC	7
XX		XX	
CC	The invention comprises a crystallisable composition containing a	SQ	Sequence 2288 AA;
CC	carboxyltransferase (CT) domain of acetyl-coenzyme A (acetyl-CoA)		
CC	carboxylase. The composition of the invention is useful for treating:		
CC	metabolic syndrome, diabetes, obesity, cardiovascular disease,		
CC	atherosclerosis, depression, cancer, hyperlipidaemia, dyslipidaemia,		
CC	hypertension, hyperuricaemia and renal disfunction. The present amino		
CC	acid sequence represents a mouse acetyl-CoA carboxylase of the invention.		
XX			
SQ	Sequence 2110 AA;		
	Query Match 73.5%; Score 36; DB 8; Length 2110;		
	Best Local Similarity 70.0%; Pred. No. 1.2e+03;		
	Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;		
QY	1 ESVREQVMKY 10		
	:		
Db	1254 ESVRSMVMRY 1263		
	RESULT 37		
AAB86033			
ID	AAB86033 standard; protein; 2288 AA.		
XX			
AC	AAB86033;		
XX			
DT	13-JUL-2001 (first entry)		
XX			
DE	Bovine acetyl-coenzyme A carboxylase-alpha protein fragment.		
XX			
KW	Acetyl-coenzyme A carboxylase-alpha; acetyl-coA carboxylase alpha;		
KW	bovine; milk gland-specific promoter; Accalpha; milk production; sheep;		
KW	goat; fat content; genotyping; lactation.		
XX			
OS	Bos taurus.		
XX			
PN	DE19946173-A1.		
XX			
PD	05-APR-2001.		
XX			
PF	20-SEP-1999; 99DE-01046173.		
XX			
PR	20-SEP-1999; 99DE-01046173.		
XX			

Query Match	73.5%;	Score	36;	DB	4;	Length	2288;
Best Local Similarity	70.0%;	Pred. No.	1.3e+03;				
Matches	7;	Conservative	1;	Mismatches	2;	Indels	0;
QY	1 ESVREQVMKY 10						
	:						
Db	1432 ESVRSMVMRY 1441						
	RESULT 38						
AAR05707							
ID	AAR05707 standard; protein; 2324 AA.						
XX							
AC	AAR05707;						
XX							
DT	25-MAR-2003 (revised)						
DT	16-AUG-1990 (first entry)						
XX							
DE	Acetyl-CoA-carboxylase.						
XX							
KW	Acetyl-CoA-carboxylase; Co2 ligase; malonyl-CoA; biotin;						
KW	biotin-carboxylase; carboxyl-transferase; acetyl-CoA; ATP; ds.						
XX							
OS	Gallus sp.						
XX							
PN	JP02057179-A.						
XX							
PD	26-FEB-1990.						
XX							
PF	24-AUG-1988; 88JP-00208170.						
XX							
PR	24-AUG-1988; 88JP-00208170.						
XX							
PA	(MEIP) MEIJI MILK PROD CO LTD.						
XX							
DR	WPI; 1990-104915/14.						
DR	N-PSDB; AAQ04013.						
XX							
PT	Acetyl-coa-carboxylase - derived from chicken liver, has known nucleotide						
PT	and amino acid sequence.						

XX PS Disclosure; Fig 2; 14pp; Japanese.

XX CC Gene derived from chicken liver. (Updated on 25-MAR-2003 to correct PA field.)

XX CC

XX SQ Sequence 2324 AA;

Query Match 73.5%; Score 36; DB 2; Length 2324;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||||| |:|

Db 1467 ESVRSMVMRY 1476

RESULT 39

ABP70270

ID ABP70270 standard; protein; 2346 AA.

XX AC ABP70270;

XX DT 07-APR-2003 (first entry)

XX DE Amino acid sequence of human acetyl coenzyme A-carboxylase alpha.

XX KW Acetyl coenzyme A-carboxylase alpha; ACC-alpha; BRCA1; breast cancer; ovarian cancer.

XX OS Homo sapiens.

XX PN WO2002100897-A2.

XX PD 19-DEC-2002.

XX PF 12-JUN-2002; 2002WO-FR002016.

PR 13-JUN-2001; 2001FR-00007740.

PR 05-MAR-2002; 2002FR-00002789.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.

XX PI Dalla Venezia N, Magnard C, Lenoir G, Sinilnikova-Erard O;

XX DR WPI; 2003-148772/14.

DR N-PSDB; AB223503.

XX PT New molecular complex of BRCA1 protein and acetyl coenzyme A-carboxylase alpha, useful in screening agents suitable for treatment, prevention or diagnosis of cancer.

XX PS Disclosure; Page 55-62; 62pp; French.

XX CC The present sequence represents human acetyl coenzyme A-carboxylase alpha (ACC-alpha). The ACC-alpha protein was used in the course of the invention. The specification describes a molecular complex comprising a polypeptide having amino acids 1640-1663 of the human BRCA1 protein (or similar sequence from some other animal species), and a polypeptide that is a part of the acetyl coenzyme A-carboxylase alpha (ACC-alpha) protein able to bind the BRCA1 protein. The complex is implicated in predisposition to cancer of breast and ovary. It is used to screen for compounds that modulate interaction between BRCA1 and ACC-alpha, which are potentially useful for treatment, prevention and diagnosis of cancer, and to identify endogenous ligands. Modulated formation of the complex can be used for diagnosis of cancer. Antibodies directed against specific parts of human ACC-alpha are useful for localization of the complex in cells

XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||||| |:|

Db 1490 ESVRSMVMRY 1499

RESULT 40

ABP59197

ID ABP59197 standard; protein; 2346 AA.

XX AC ABP59197;

XX DT 01-MAY-2003 (first entry)

XX DE Human acetyl-Coenzyme A-carboxylase-alpha variant #1.

XX KW Human; enzyme; acetyl-Coenzyme A-carboxylase-alpha; ACC-alpha; cancer; breast; ovary.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1405 /note= "Leu substituted with Val"

FT

XX WO2002100896-A2.

PN 19-DEC-2002.

XX PD 12-JUN-2002; 2002WO-FR002015.

XX PF 13-JUN-2001; 2001FR-00007740.

PR 05-MAR-2002; 2002FR-00002788.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.

XX PI Dalla Venezia NL, Magnard CM, Lenoir GM, Sinilnikova-Erard O;

XX DR WPI; 2003-175165/17.

XX PT In vitro diagnosis of cancer, particularly breast and ovarian cancer, or susceptibility, comprises detecting alterations in the acetyl coenzyme A-carboxylase alpha gene or protein expression.

XX PS Example 2; Page; 56pp; French.

XX CC The present sequence is a variant sequence for human acetyl-Coenzyme A-carboxylase-alpha (ACC-alpha). This sequence can be used for in vitro diagnosis of cancer (or of an increased risk of developing it), by detecting ACC-alpha gene mutations or polymorphisms, or altered ACC-alpha protein expression, relative to a control population. The method is particularly used to diagnose cancer, especially of breast or ovary, or for assessing the risk of developing such cancers. Note: The present sequence was not shown in the specification, but was derived from information given

XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||||| |:|

Db 1490 ESVRSMVMRY 1499

RESULT 41

ABP59195

ID ABP59195 standard; protein; 2346 AA.

XX AC ABP59195;
XX DT 01-MAY-2003 (first entry)
XX DE Human acetyl-Coenzyme A-carboxylase-alpha.
XX KW Human; enzyme; acetyl-Coenzyme A-carboxylase-alpha; ACC-alpha; cancer;
XX KW breast; ovary.
XX OS Homo sapiens.
XX PN WO2002100896-A2.
XX PD 19-DEC-2002.
XX PF 12-JUN-2002; 2002WO-FR002015.
XX PR 13-JUN-2001; 2001FR-00007740.
XX PR 05-MAR-2002; 2002FR-00002788.
XX PA (CNRS) CNRS CENT NAT RECH SCI.
XX PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
XX PI Dalla Venezia NL, Magnard CM, Lenoir GM, Sinilnikova-Erard O;
XX DR N-PSDB; ABZ79442.
XX PT In vitro diagnosis of cancer, particularly breast and ovarian cancer, or
XX PT susceptibility, comprises detecting alterations in the acetyl coenzyme A-
XX PT carboxylase alpha gene or protein expression.
XX PS Disclosure; Page 49-56; 56pp; French.
XX CC The present sequence is the protein sequence for human acetyl-Coenzyme A-
XX CC carboxylase-alpha (ACC-alpha). This sequence can be used for in vitro
XX CC diagnosis of cancer (or of an increased risk of developing it), by
XX CC detecting ACC-alpha gene mutations or polymorphisms, or altered ACC-alpha
XX CC protein expression, relative to a control population. The method is
XX CC particularly used to diagnose cancer, especially of breast or ovary, or
XX CC for assessing the risk of developing such cancers
XX CC
XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 1490 ESVRSMVMRY 1499

RESULT 42
ABP59198
ID ABP59198 standard; protein; 2346 AA.
XX AC ABP59198;
XX DT 01-MAY-2003 (first entry)
XX DE Human acetyl-Coenzyme A-carboxylase-alpha variant #2.
XX KW Human; enzyme; acetyl-Coenzyme A-carboxylase-alpha; ACC-alpha; cancer;
XX KW breast; ovary.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 2271
XX FT /note= "Ala substituted with Val"

PN WO2002100896-A2.
XX PD 19-DEC-2002.
XX PF 12-JUN-2002; 2002WO-FR002015.
XX PR 13-JUN-2001; 2001FR-00007740.
XX PR 05-MAR-2002; 2002FR-00002788.
XX PA (CNRS) CNRS CENT NAT RECH SCI.
XX PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
XX PI Dalla Venezia NL, Magnard CM, Lenoir GM, Sinilnikova-Erard O;
XX DR WPI; 2003-175165/17.
XX PT In vitro diagnosis of cancer, particularly breast and ovarian cancer, or
XX PT susceptibility, comprises detecting alterations in the acetyl coenzyme A-
XX PT carboxylase alpha gene or protein expression.
XX PS Example 2; Page; 56pp; French.
XX CC The present sequence is a variant sequence for human acetyl-Coenzyme A-
XX CC carboxylase-alpha (ACC-alpha). This sequence can be used for in vitro
XX CC diagnosis of cancer (or of an increased risk of developing it), by
XX CC detecting ACC-alpha gene mutations or polymorphisms, or altered ACC-alpha
XX CC protein expression, relative to a control population. The method is
XX CC particularly used to diagnose cancer, especially of breast or ovary, or
XX CC for assessing the risk of developing such cancers. Note: The present
XX CC sequence was not shown in the specification, but was derived from
XX CC information given
XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 1490 ESVRSMVMRY 1499

RESULT 43
ADQ90744
ID ADQ90744 standard; protein; 2346 AA.
XX AC ADQ90744;
XX DT 21-OCT-2004 (first entry)
XX DE Human acetyl-coenzyme A carboxylase - ACC1.
XX KW human; carboxyltransferase domain; CT domain;
XX KW acetyl-coenzyme A carboxylase; acetyl-CoA carboxylase;
XX KW metabolic syndrome; diabetes; obesity; cardiovascular disease;
XX KW atherosclerosis; depression; cancer; hyperlipidaemia; dyslipidaemia;
XX KW hypertension; hyperuricaemia; renal dysfunction;
XX KW crystallisable composition; enzyme; ACC1.
XX OS Homo sapiens.
XX PN WO2004063715-A2.
XX PD 29-JUL-2004.
XX PF 09-JAN-2004; 2004WO-US000585.
XX PR 10-JAN-2003; 2003US-0439383P.
XX PR 31-MAR-2003; 2003US-0459464P.
XX PR 31-JUL-2003; 2003US-0491640P.
XX PR 27-OCT-2003; 2003US-0514636P.
XX PR 09-JAN-2004; 2004US-00514636.

XX (UYCO) UNIV COLUMBIA NEW YORK.
PA
XX Tong L, Zhang H, Yang Z;
PI
XX WPI; 2004-571486/55.
DR
XX New crystallizable composition comprising a carboxyltransferase domain of
PT acetyl-coenzyme A (CoA) carboxylase, for treating e.g. diabetes, obesity,
PT cardiovascular disease, atherosclerosis, or cancer.
XX
PS Example 4; SEQ ID NO 7; 195pp; English.
XX
CC The invention comprises a crystallisable composition containing a
CC carboxyltransferase (CT) domain of acetyl-coenzyme A (acetyl-CoA)
CC carboxylase. The composition of the invention is useful for treating:
CC metabolic syndrome, diabetes, obesity, cardiovascular disease,
CC atherosclerosis, depression, cancer, hyperlipidaemia, dyslipidaemia,
CC hypertension, hyperuricaemia and renal dysfunction. The present amino
CC acid sequence represents a human acetyl-CoA carboxylase of the invention.
XX
SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 8; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||||| ||:|
Db 1490 ESVRSMVMRY 1499

RESULT 44
ADA33386
ID ADA33386 standard; protein; 232 AA.
XX
AC ADA33386;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #547.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR N-PSDB; ADA29260.
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 4673; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 232 AA;

Query Match 71.4%; Score 35; DB 6; Length 232;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||: ||:|
Db 8 ESLSEQIVKY 17

RESULT 45
ABR53195
ID ABR53195 standard; protein; 810 AA.
XX
AC ABR53195;
XX
DT 20-JUN-2003 (first entry)
XX
DE Protein sequence #SEQ ID 1255.
XX
KW Multiprotein complex; eukaryote; drug target; diagnosis.
XX
OS Saccharomyces cerevisiae.
XX
PN EP1258494-A1.
XX
PD 20-NOV-2002.
XX
PF 20-DEC-2001; 2001EP-00130253.
XX
PR 15-MAY-2001; 2001EP-00111774.
XX
PA (CELL-) CELLZOME AG.
XX
PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
PI Marzioch M, Schultz JD, Superti-Furga GD;
XX
DR WPI; 2003-250078/25.
DR N-PSDB; ACC61237.
XX
XX New isolated protein complexes useful for diagnosing a disease or
PT disorder, or as a target for an active agent of a pharmaceutical,
PT preferably a drug target in the treatment or prevention of disease or
PT disorder.
XX
PS Disclosure; SEQ ID NO 1255; 17pp + Sequence Listing; English.
XX
CC The invention relates to multiprotein complexes from eukaryotes. Proteins
CC of the invention and DNA sequences encoding them are given in records
CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
CC obtainable by using a protein as a bait and isolating the set of proteins
CC which is attached thereto from cells. Such protein complexes may comprise
CC up to 30 distinct proteins. Protein complexes of the invention are useful
CC for diagnosing a disease or disorder, or as a target for an active agent
CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: The sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM
XX
SQ Sequence 810 AA;

Query Match 71.4%; Score 35; DB 6; Length 810;
Best Local Similarity 70.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | : | |
Db 13 ESVEEAVLKY 22

RESULT 46
ADK62816
ID ADK62816 standard; protein; 810 AA.
XX
AC ADK62816;
XX
DT 06-MAY-2004 (first entry)
XX
DE Disease treating protein complex-derived protein #534.
XX
KW protein complex; drug target; diagnosis.
XX
OS Unidentified.
XX
PN EP1338608-A2.
XX
PD 27-AUG-2003.
XX
PF 20-DEC-2002; 2002EP-00102902.
XX
PR 20-DEC-2001; 2001EP-00130253.
XX
PA (CELL-) CELLZOME AG.
XX

PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
PI Michon A, Leutwein C, Rick J;
XX
DR WPI; 2003-638460/61.
DR N-PSDB; ADK62817.
XX

PT New proteins and protein complexes from eukaryotes, useful as targets in
PT drug screening, or in diagnosing or screening for the presence of a
PT disease or disorder, or a predisposition for developing a disease or
PT disorder in a subject.
XX
PS Disclosure; SEQ ID NO 1067; 13pp; English.
XX

CC The invention relates to novel protein complexes comprising a first and a
CC second protein, or its derivative, fragment, homologue or variant. The
CC proteins are selected from given protein complexes, which are not defined
CC in the specification. The variants are encoded by nucleic acids that
CC hybridize to the nucleic acids encoding the proteins under low stringency
CC conditions. The protein complexes are useful as targets for an active
CC agent of a pharmaceutical. These protein complexes are particularly
CC useful as drugs targets for the treatment or preventing of a disease or
CC disorder. The complexes and methods above are useful in diagnosing or
CC screening for the presence of a disease or disorder or a predisposition
CC for developing a disease or disorder in a subject. These are also useful
CC in screening for a drug for treatment or prevention of a disease or
CC disorder. The molecule that modulates the amount, activity or protein
CC components of the complex is useful for the manufacture of a medicament
CC for the treatment or prevention of a disease or disorder. This sequence
CC corresponds to a protein of the invention. (Note: the sequence data for
CC this patent did not form part of the printed specification but was
CC obtained from the EPO in electronic format).
XX

SQ Sequence 810 AA;

Query Match 71.4%; Score 35; DB 7; Length 810;
Best Local Similarity 70.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | : | |
Db 13 ESVEEAVLKY 22

RESULT 47
ADO16901
ID ADO16901 standard; protein; 1245 AA.
XX
AC ADO16901;
XX

DT 12-AUG-2004 (first entry)
XX
DE Tobacco budworm voltage gated calcium channel subunit protein SeqID 14.
XX
KW voltage-gated calcium channel; pesticide; agrochemical; veterinary;
KW pharmaceutical; insecticide; nematocide; molluscide; helminthicide;
KW acaricide; biocide; crop protection agent; antiparasitic;
KW tobacco budworm.
XX

OS Heliothis virescens.
XX
PN WO2004044553-A2.
XX

PD 27-MAY-2004.
XX

PF 10-NOV-2003; 2003WO-US036001.
XX

PR 08-NOV-2002; 2002US-0425077P.
XX

PA (FMCC) FMC CORP.
XX

PI Wu S, Hayashi JH, Kinne LP, Dierks PM;
XX

DR WPI; 2004-411759/38.
DR N-PSDB; ADO16900.
XX

PT New Lepidoptera calcium channel subunits, useful as active agents in
PT agrochemical, veterinary or pharmaceutical fields, e.g. as insecticide,
PT acaricide, or for treating infection, damage or discomfort caused by
PT parasitic organisms.

XX Claim 1; SEQ ID NO 14; 144pp; English.

PS This invention relates to novel isolated nucleic acids from Lepidoptera
XX that each encode a voltage-gated calcium channel subunit protein.
CC Specifically, it refers to methods and screening assays to identify
CC compounds that can target and modulate these channels and as such that
CC have potential activity as pesticides. Accordingly, the present invention
CC provides active agents that have applications in the agrochemical,
CC veterinary and pharmaceutical fields, and can be used as an insecticide,
CC nematocide, molluscide, helminthicide, acaricide or other types of
CC pesticides or biocides. In addition, they can be applied systemically (as
CC additives) to the soil or habitat of the organism to act as crop
CC protection agents or as a pesticide for household use to prevent or treat
CC infection and/ or discomfort caused by parasitic organisms such as
CC parasitic arthropods, nematodes or helminths. This polypeptide sequence
CC is a tobacco budworm voltage-gated calcium channel protein subunit of the
CC invention.

SQ Sequence 1245 AA;

Query Match 71.4%; Score 35; DB 8; Length 1245;
Best Local Similarity 70.0%; Pred. No. 1e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | : | |
Db 415 EEVREQVVHY 424

RESULT 48

AAY22382

ID AAY22382 standard; peptide; 15 AA.

XX

AC AAY22382;

XX

DT 27-SEP-1999 (first entry)

XX TPO receptor binding peptide sequence, SEQ ID NO. 33.
DE
XX
XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
XX US5932546-A.
PN
XX
XX 03-AUG-1999.
PD
XX
XX 04-OCT-1996; 96US-00726464.
PF
XX
XX 04-OCT-1996; 96US-00726464.
PR
XX
XX (GLAX) GLAXO WELLCOME INC.
PA
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirila SE, Johnson SS;
XX
XX WPI; 1999-457122/38.
DR
XX
XX New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
PT
XX
XX Disclosure; Col 13-14; 36pp; English.
PS
XX
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;
PS
XX
XX Query Match 69.4%; Score 34; DB 2; Length 15;
CC Best Local Similarity 70.0%; Pred. No. 13;
CC Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC
CC 1 ESVREQVMKY 10
QY :|||||||
DB 4 QSVREQVMWAY 13
XX
XX
XX RESULT 49
AAY22374
ID AAY22374 standard; peptide; 15 AA.
XX
XX AC AAY22374;
AC
XX
XX 27-SEP-1999 (first entry)
DT
XX
XX TPO receptor binding peptide sequence, SEQ ID NO. 25.
DE
XX
XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
XX Synthetic.
OS
XX
XX US5932546-A.
PN
XX
XX 03-AUG-1999.
PD
XX
XX 04-OCT-1996; 96US-00726464.
PF

XX 04-OCT-1996; 96US-00726464.
PR (GLAX) GLAXO WELLCOME INC.
XX
XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirila SE, Johnson SS;
PI
XX
XX WPI; 1999-457122/38.
DR
XX
XX New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
PT
XX
XX Disclosure; Col 13-14; 36pp; English.
PS
XX
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration
XX
SQ Sequence 15 AA;
PS
XX
XX Query Match 69.4%; Score 34; DB 2; Length 15;
CC Best Local Similarity 77.8%; Pred. No. 13;
CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC
CC 2 SVREQVMKY 10
QY :|||||||
DB 5 SVREQVMMAF 13
XX
XX
XX RESULT 50
AAG92828
ID AAG92828 standard; protein; 212 AA.
XX
XX AC AAG92828;
AC
XX
XX 26-SEP-2001 (first entry)
DT
XX
XX C glutamicum protein fragment SEQ ID NO: 6582.
DE
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
KW
XX
XX Corynebacterium glutamicum.
OS
XX
XX EP1108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-00127688.
PF
XX
XX 16-DEC-1999; 99JP-00377484.
PR
XX
XX 07-APR-2000; 2000JP-00159162.
PR
XX
XX 03-AUG-2000; 2000JP-00280988.
PR
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI
XX
XX WPI; 2001-376931/40.
DR
XX
XX N-PSDB; AAH68047.
DR
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT

PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 6582; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 212 AA;
Query Match 69.4%; Score 34; DB 4; Length 212;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
|:|||||:
Db 128 EEIREQVLLY 137
Search completed: May 12, 2006, 10:36:36
Job time : 102.75 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-38
Perfect score: 49
Sequence: 1 ESVREQVMKY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	81.6	1164	2	G71827	hypothetical prote
2	38	77.6	291	2	S73600	probable GTP-bindi
3	37	75.5	429	2	A47305	translation initia
4	36	73.5	151	2	S54048	ribosomal protein
5	36	73.5	160	2	F64816	ybiA protein - Esc
6	36	73.5	1234	2	T30254	jumonji protein -
7	36	73.5	2324	1	A29924	acetyl-CoA carboxy
8	36	73.5	2345	1	A35578	acetyl-CoA carboxy
9	36	73.5	2346	2	I38928	acetyl-CoA carboxy
10	35	71.4	486	2	G72358	conserved hypothet
11	35	71.4	1279	1	S45907	myb-related protei
12	35	71.4	1279	1	E69681	peptide synthetase
13	34	69.4	212	2	G82176	probable outer mem
14	34	69.4	272	2	B90544	hypothetical prote
15	34	69.4	357	2	AD1062	protein kinase [im
16	34	69.4	455	2	E90514	hypothetical prote
17	34	69.4	595	2	T17590	probable glutamine
18	34	69.4	786	2	S71091	acetyl-CoA carboxy
19	34	69.4	825	2	S54465	YTA12 protein prec
20	34	69.4	916	2	C38992	cadherin 4 precurs
21	34	69.4	2339	2	S41121	acetyl-CoA carboxy
22	33	67.3	179	2	B90408	hypothetical prote
23	33	67.3	243	2	T25190	hypothetical prote
24	33	67.3	278	2	T03942	rRNA N-glycosidase
25	33	67.3	282	2	A84341	hypothetical prote
26	33	67.3	298	2	G82581	GTP binding protei
27	33	67.3	355	2	H69486	conserved hypothet
28	33	67.3	358	2	T43289	fructose-bisphosph
29	33	67.3	358	2	T39798	fructose-bisphosph

30	33	67.3	392	2	AE2092	hypothetical prote
31	33	67.3	431	2	T12450	hypothetical prote
32	33	67.3	490	2	F86841	iron-binding oxida
33	33	67.3	514	2	E84667	probable cytochrom
34	33	67.3	570	2	A45249	alpha-glucosidase
35	33	67.3	572	2	T41371	hypothetical protei
36	33	67.3	595	1	A48077	myb-related protei
37	33	67.3	1153	2	A97179	ATP-dependent exon
38	33	67.3	1202	2	T23429	hypothetical prote
39	33	67.3	1873	2	A30063	dihydropyridine re
40	32	65.3	92	2	S69116	fibrinogen gamma c
41	32	65.3	138	2	T27724	hypothetical prote
42	32	65.3	217	2	E97310	uncharacterized co
43	32	65.3	222	2	H64321	hypothetical prote
44	32	65.3	301	2	S45880	SCO1 protein homol
45	32	65.3	314	2	S00909	replication initia
46	32	65.3	314	2	A30480	repJ protein - Sta
47	32	65.3	316	2	T49224	SRG1-like protein
48	32	65.3	319	2	T04309	acidic ribosomal p
49	32	65.3	319	2	T03944	acidic ribosomal p
50	32	65.3	350	2	T22701	hypothetical prote
51	32	65.3	376	2	S49801	OGG1 protein - yea
52	32	65.3	403	2	E82810	methionine adenosy
53	32	65.3	403	2	D96961	carboxyl-terminal
54	32	65.3	410	2	G72586	hypothetical prote
55	32	65.3	444	2	S05313	fibrinogen gamma-B
56	32	65.3	445	2	E83805	hypothetical prote
57	32	65.3	445	2	H97075	4-aminobutyrate am
58	32	65.3	489	1	D64311	fumarate reductase
59	32	65.3	501	2	G72368	comM protein - The
60	32	65.3	530	2	T46648	zinc finger protei
61	32	65.3	550	2	H84830	hypothetical prote
62	32	65.3	570	1	S50933	myb-related protei
63	32	65.3	678	2	T50256	probable vacuolar
64	32	65.3	721	2	H82528	L-ascorbate oxidas
65	32	65.3	759	2	B81101	ribonucleoside-dip
66	32	65.3	782	2	AD0119	conserved hypothet
67	32	65.3	782	2	AF0179	conserved hypothet
68	32	65.3	793	2	D38992	cadherin 8 - human
69	32	65.3	845	2	T00071	hypothetical prote
70	32	65.3	930	2	AC2412	preprotein translo
71	32	65.3	940	2	A82329	excinuclease ABC c
72	32	65.3	1059	2	T22545	hypothetical prote
73	32	65.3	2156	1	RRVUNE	genome polypeptide
74	31	63.3	91	2	G75094	lsu ribosomal prot
75	31	63.3	94	2	E71071	probable ribosomal
76	31	63.3	96	2	F97355	uncharacterized co
77	31	63.3	120	2	E81229	30S ribosomal prot
78	31	63.3	127	2	C71713	hypothetical prote
79	31	63.3	130	2	A48892	abscisic acid-indu
80	31	63.3	139	2	E97709	hypothetical prote
81	31	63.3	166	2	G83399	hypothetical prote
82	31	63.3	188	2	G96652	protein F23N19.15
83	31	63.3	189	2	C81428	peptidyl-prolyl ci
84	31	63.3	204	1	MNNZHS	nonstructural prot
85	31	63.3	204	1	MNNZSV	nonstructural prot
86	31	63.3	206	2	A98106	hypothetical prote
87	31	63.3	206	2	F95241	hydrolase, haloaci
88	31	63.3	259	2	S36410	translation initia
89	31	63.3	268	2	E82472	hypothetical prote
90	31	63.3	283	2	A96655	hypothetical prote
91	31	63.3	291	2	A75096	UTP-glucose-1-phos
92	31	63.3	292	2	B90454	dehydrogenase [imp
93	31	63.3	292	2	AH1354	ABC transporter (A
94	31	63.3	294	2	S47535	homeodomain-contai
95	31	63.3	295	2	S29045	estrone sulfotrans
96	31	63.3	311	1	RQSAD2	repD protein - Sta
97	31	63.3	314	2	A26473	replication protei
98	31	63.3	316	1	A39763	aldehyde reductase
99	31	63.3	330	2	B97886	hypothetical prote
100	31	63.3	341	2	B86598	flagellar M-ring p
101	31	63.3	341	2	D72025	flagellar m-ring p
102	31	63.3	386	2	C70505	probable acyl-coa

103 31 63.3 415 2 AI0361 conserved hypothet
104 31 63.3 430 2 S62500 actin-related prot
105 31 63.3 443 2 JC4067 hypothetical 50.1k
106 31 63.3 488 2 A95926 probable argininos
107 31 63.3 590 2 A82350 conserved hypothet
108 31 63.3 593 2 AD2328 hypothetical prote
109 31 63.3 632 2 H69590 asparagine synthet
110 31 63.3 765 2 E88924 protein R02C2.3 [i
111 31 63.3 856 2 E70875 hypothetical prote
112 31 63.3 924 2 C86725 chromosome segrega
113 31 63.3 925 2 T33732 probable excinucle
114 31 63.3 932 2 C70191 penicillin-binding
115 31 63.3 1043 2 AH1906 hypothetical prote
116 31 63.3 1076 2 T38122 hypothetical prote
117 31 63.3 1116 2 T14598 polyprotein - slim
118 31 63.3 1164 2 T24806 hypothetical prote
119 31 63.3 2223 2 A47447 calcium channel pr
120 30 61.2 87 2 T10462 hypothetical prote
121 30 61.2 110 2 S74013 calcium channel pr
122 30 61.2 115 2 E36811 hypothetical prote
123 30 61.2 118 2 AH0028 hypothetical prote
124 30 61.2 134 2 I47010 hypothetical prote
125 30 61.2 140 2 F75335 30S ribosomal prot
126 30 61.2 144 2 AI2573 gastrin-releasing
127 30 61.2 151 1 S25374 probable transposa
128 30 61.2 151 1 S26296 hypothetical prote
129 30 61.2 154 2 D69137 RNA polymerase sig
130 30 61.2 194 2 G82091 probable phage tai
131 30 61.2 201 2 AI0928 probable phage tai
132 30 61.2 201 2 AI1035 hypothetical prote
133 30 61.2 210 2 F96551 probable exported
134 30 61.2 226 2 AC0176 hypothetical prote
135 30 61.2 229 2 A69762 peripheral membran
136 30 61.2 241 1 BVECPU phosphate transpor
137 30 61.2 241 2 AF0956 hypothetical prote
138 30 61.2 241 2 D86057 hypothetical prote
139 30 61.2 241 2 D91211 hypothetical prote
140 30 61.2 249 2 F95957 probable transcrip
141 30 61.2 249 2 E84072 hypothetical prote
142 30 61.2 256 2 A75005 hypothetical prote
143 30 61.2 256 2 H71157 hypothetical prote
144 30 61.2 262 2 G82280 vibriobactin-speci
145 30 61.2 273 2 F71318 conserved hypothet
146 30 61.2 284 2 H86183 protein T7A14.13 [i
147 30 61.2 295 2 F59102 probable UTP-glucos
148 30 61.2 303 1 CPCYB carboxypeptidase B
149 30 61.2 305 2 T34177 hypothetical prote
150 30 61.2 321 1 R5UBP0 acidic ribosomal p
151 30 61.2 335 2 T25110 hypothetical prote
152 30 61.2 346 2 F87844 protein T22C1.6 [i
153 30 61.2 367 2 H82437 2-aminoethylphosph
154 30 61.2 382 2 C87473 acyl-CoA dehydroge
155 30 61.2 387 2 S74022 hypothetical prote
156 30 61.2 396 2 F84326 GTP-binding protei
157 30 61.2 406 1 S48220 serine-type D-Ala-
158 30 61.2 409 2 T14611 alanine transamina
159 30 61.2 430 2 G86870 histidine-tRNA lig
160 30 61.2 432 2 E89954 conserved hypothet
161 30 61.2 448 2 T18710 hypothetical prote
162 30 61.2 459 2 JC7931 cellobiohydrolase
163 30 61.2 463 2 S06309 vicilin precursor
164 30 61.2 463 2 A27288 vicilin precursor
165 30 61.2 466 2 E75201 pyridoxal phosphat
166 30 61.2 470 2 F72238 conserved hypothet
167 30 61.2 471 2 A45616 antigenic protein
168 30 61.2 481 2 E90203 prolyl -tRNA synth
169 30 61.2 483 2 AB0072 tagaturonate reduc
170 30 61.2 489 2 C84401 aldehyde dehydroge
171 30 61.2 497 2 T14609 alanine transamina
172 30 61.2 501 2 T37747 hypothetical prote
173 30 61.2 506 1 PWEGA H+-transporting tw
174 30 61.2 518 2 B84514 probable cytochrom
175 30 61.2 532 2 E69343 2-oxoacid-ferredox

176 30 61.2 553 2 AC2931 hypothetical prote
177 30 61.2 603 2 H69121 hypothetical prote
178 30 61.2 619 2 S30780 hexose metabolism-
179 30 61.2 626 2 D88601 protein Y49E10.11
180 30 61.2 630 2 C98351 oligopeptide-bindi
181 30 61.2 657 2 G86590 exinuclease ABC su
182 30 61.2 657 2 B72034 exinuclease ABC,
183 30 61.2 698 2 G89787 hypothetical prote
184 30 61.2 727 2 T26096 hypothetical prote
185 30 61.2 752 2 G96510 hypothetical prote
186 30 61.2 759 2 G81841 ribonucleoside-dip
187 30 61.2 769 1 ISBYT1 DNA topoisomerase
188 30 61.2 774 1 I40421 endopeptidase La (
189 30 61.2 774 2 B84031 ATP-dependent prot
190 30 61.2 779 2 T05990 hypothetical prote
191 30 61.2 783 2 I50116 N-cadherin precurs
192 30 61.2 783 2 AH3592 ribonuclease R (EC
193 30 61.2 786 2 F84608 hypothetical prote
194 30 61.2 790 2 G02678 cadherin-14 - huma
195 30 61.2 791 2 D82901 ATP-dependent prot
196 30 61.2 796 2 I49556 cadherin-11 - mous
197 30 61.2 798 2 G83420 Lon proteinase PA1
198 30 61.2 799 2 C87492 ATP-dependent prot
199 30 61.2 805 2 A75014 hypothetical prote
200 30 61.2 807 2 T00990 hypothetical prote
201 30 61.2 810 1 JC6045 endopeptidase La (
202 30 61.2 820 2 H81106 ATP-dependent prot
203 30 61.2 820 2 H81908 probable endopepti
204 30 61.2 858 2 D71711 endopeptidase Clp
205 30 61.2 877 1 IJBOCN N-cadherin precurs
206 30 61.2 881 1 B71316 endopeptidase La (
207 30 61.2 889 2 T20123 hypothetical prote
208 30 61.2 896 2 S76064 hypothetical prote
209 30 61.2 896 2 S59990 phycobilisome anch
210 30 61.2 905 1 IJXLC1 N-cadherin 1 precu
211 30 61.2 905 1 RGBYS5 regulatory protein
212 30 61.2 906 1 IJHUCN cadherin 2 precurs
213 30 61.2 906 1 IJM5CN N-cadherin precurs
214 30 61.2 906 1 IJXLC2 N-cadherin 2 precu
215 30 61.2 906 2 T28034 hypothetical prote
216 30 61.2 912 1 IJCHCN N-cadherin precurs
217 30 61.2 913 1 A47543 R-cadherin precurs
218 30 61.2 913 1 IJCHCR R-cadherin precurs
219 30 61.2 976 2 S45738 pleiotropic drug r
220 30 61.2 1007 2 S48535 rho-type GTPase-ac
221 30 61.2 1080 2 A35088 phycobilisome link
222 30 61.2 1123 2 D96756 receptor-like prot
223 30 61.2 1132 2 AD1809 phycobilisome core
224 30 61.2 1286 2 A42150 P-glycoprotein pgp
225 30 61.2 1314 2 S19488 probable membrane
226 30 61.2 1802 2 H88444 protein C2E6.12 [i
227 30 61.2 1811 2 T00035 nonstructural poly
228 30 61.2 2151 1 S16449 genome polypoteins
229 30 61.2 2160 2 T20241 hypothetical prote
230 30 61.2 2206 2 JC5280 voltage-dependent
231 30 61.2 2523 2 T18477 hypothetical prote
232 29 59.2 76 2 AG0683 hypothetical prote
233 29 59.2 82 2 AB2351 hypothetical prote
234 29 59.2 83 1 W8BPT3 hypothetical prote
235 29 59.2 97 2 S06782 gene 18.7 protein
236 29 59.2 98 2 AC1033 tryptophan synthas
237 29 59.2 134 2 E57233 hypothetical prote
238 29 59.2 134 2 JC4226 complexin II - hum
239 29 59.2 134 2 D57233 complexin II - rat
240 29 59.2 134 2 D57233 complexin II - mou
241 29 59.2 136 2 B97881 hypothetical prote
242 29 59.2 156 2 C81025 H+-transporting tw
243 29 59.2 157 2 AG3560 transcription regu
244 29 59.2 158 2 A84069 hypothetical prote
245 29 59.2 165 2 AC0451 conserved hypothet
246 29 59.2 172 2 B64320 acetolactate synth
247 29 59.2 180 2 F69408 phosphoribosylamin
248 29 59.2 181 2 D81285 hypothetical prote

249	29	59.2	181	2	T44572	ribosomal protein	322	29	59.2	531	2	T25537	hypothetical prote
250	29	59.2	181	2	D75052	ribosomal protein	323	29	59.2	532	2	B64100	purH bifunctional
251	29	59.2	181	2	A71108	ribosomal protein	324	29	59.2	545	2	F84533	Mutator-like trans
252	29	59.2	196	2	AF0699	probable pathogeni	325	29	59.2	564	2	AD1550	heat shock protein
253	29	59.2	197	2	F69359	conserved hypoteth	326	29	59.2	571	2	C84038	DNA-dependent DNA
254	29	59.2	198	2	S73275	hypothetical prote	327	29	59.2	588	2	C95252	L-fucose isomerase
255	29	59.2	224	2	H84829	hypothetical prote	328	29	59.2	588	2	A99717	L-fucose isomerase
256	29	59.2	233	2	C86849	regulatory protein	329	29	59.2	598	2	T35513	probable long-chai
257	29	59.2	235	2	E72217	hypothetical prote	330	29	59.2	611	2	D72275	conserved hypoteth
258	29	59.2	248	2	T23708	hypothetical prote	331	29	59.2	648	2	H84587	probable WD-40 rep
259	29	59.2	248	2	C84741	hypothetical prote	332	29	59.2	650	2	H81708	probable bacteriop
260	29	59.2	254	2	B72366	conserved hypoteth	333	29	59.2	650	2	AG0733	lysostaphin (EC 3.
261	29	59.2	261	2	T46138	myb-like protein -	334	29	59.2	651	2	AH3276	hypothetical prote
262	29	59.2	267	2	D70376	conserved hypoteth	335	29	59.2	653	2	A84675	hypothetical prote
263	29	59.2	274	2	A11256	hypothetical prote	336	29	59.2	667	2	B96575	hypothetical prote
264	29	59.2	274	2	AE1619	hypothetical prote	337	29	59.2	677	2	F95232	immunity protein,
265	29	59.2	275	2	G89961	hypothetical prote	338	29	59.2	680	2	H98096	conserved hypoteth
266	29	59.2	287	2	G70974	hypothetical prote	339	29	59.2	686	2	C64428	hypothetical prote
267	29	59.2	287	2	A70614	probable cmaA1 pro	340	29	59.2	702	2	T03903	hypothetical prote
268	29	59.2	290	2	D81359	probable integral	341	29	59.2	711	2	S67293	probable membrane
269	29	59.2	290	2	H64242	GTP-binding protei	342	29	59.2	726	2	AG0837	TonB-dependent out
270	29	59.2	296	1	F69257	MJ1225-related pro	343	29	59.2	746	2	A75018	transport protein
271	29	59.2	310	2	T01224	probable methylene	344	29	59.2	759	2	G69258	ATP-dependent RNA
272	29	59.2	310	2	F64311	hypothetical prote	345	29	59.2	762	2	T00410	protein kinase hom
273	29	59.2	312	2	E86705	integrase-recombin	346	29	59.2	779	2	E97778	endopeptidase Ia (
274	29	59.2	317	2	F96810	hypothetical prote	347	29	59.2	784	1	A71704	ATP-dependent Ia p
275	29	59.2	323	2	G86877	glucokinase (EC 2.	348	29	59.2	786	2	G82141	ribonucleoside-dip
276	29	59.2	325	2	A82073	GTP-binding protei	349	29	59.2	786	2	A69927	ribonucleoside-dip
277	29	59.2	335	2	AC0786	nucleoid-associated	350	29	59.2	789	2	T42564	cadherin 11 precur
278	29	59.2	335	2	A64988	hypothetical 37.8	351	29	59.2	796	2	I48277	cadherin-11 - mous
279	29	59.2	335	2	H85857	nucleoid-associate	352	29	59.2	796	2	AS3584	OB-cadherin precur
280	29	59.2	335	2	F91013	nucleoid-associate	353	29	59.2	796	2	AS3584	hypothetical prote
281	29	59.2	338	2	G83508	flagellar motor sw	354	29	59.2	799	2	T48690	hypothetical prote
282	29	59.2	339	2	T22571	hypothetical prote	355	29	59.2	804	2	T18014	collagenase - Meth
283	29	59.2	341	2	F87620	conserved hypoteth	356	29	59.2	807	2	D69102	ribonuclease [impo
284	29	59.2	345	1	MFVNSY	matrix protein - S	357	29	59.2	817	2	F86742	translation initia
285	29	59.2	355	2	T02056	protein kinase MEK	358	29	59.2	864	2	A84974	sensor protein Kdp
286	29	59.2	367	2	T29752	hypothetical prote	359	29	59.2	900	2	D97351	hypothetical prote
287	29	59.2	370	2	D84464	hypothetical prote	360	29	59.2	906	2	T00039	probable helicase
288	29	59.2	377	2	A53384	polysialic acid ca	361	29	59.2	922	2	T38371	excinuclease ABC c
289	29	59.2	377	2	S60758	siaA protein - Nei	362	29	59.2	948	2	B81883	excinuclease ABC c
290	29	59.2	378	2	JH0134	creatinase (EC 3.5	363	29	59.2	948	2	AD0790	excinuclease ABC c
291	29	59.2	387	1	A56466	mitogen-activated	364	29	59.2	949	2	A81138	probable zinc-bind
292	29	59.2	390	2	B95954	probable membrane-	365	29	59.2	965	2	C82560	DNA topoisomerase
293	29	59.2	392	2	I51422	glutamine syntheta	366	29	59.2	999	2	T36021	SEC14 protein homo
294	29	59.2	392	2	S32405	MFT1 protein - yea	367	29	59.2	1003	2	T13951	hypothetical prote
295	29	59.2	392	2	T46418	hypothetical prote	368	29	59.2	1008	2	T41244	hypothetical prote
296	29	59.2	396	2	A47151	methionine adenosy	369	29	59.2	1013	2	C83771	transposase - Esch
297	29	59.2	397	2	S06114	methionine adenosy	370	29	59.2	1015	1	TQECT	fibronectin-bindin
298	29	59.2	425	2	D72345	trigger factor TM0	371	29	59.2	1018	2	A32192	RNA helicase HEL11
299	29	59.2	432	2	T05236	hypothetical prote	372	29	59.2	1032	2	A57514	beta-galactosidase
300	29	59.2	434	2	A43252	probable transcrip	373	29	59.2	1034	2	T30551	beta-galactosidase
301	29	59.2	436	2	C95222	RNA methyltransfer	374	29	59.2	1034	2	T30574	DNA-directed DNA p
302	29	59.2	442	2	B82899	conserved hypoteth	375	29	59.2	1078	2	AF3486	DNA polymerase III
303	29	59.2	444	2	H75544	hypothetical prote	376	29	59.2	1083	2	A87647	ribonucleotide red
304	29	59.2	451	2	C98086	conserved hypoteth	377	29	59.2	1084	2	T12925	hypothetical prote
305	29	59.2	452	2	AB2965	secretion protein,	378	29	59.2	1134	2	D75014	transcription acti
306	29	59.2	452	2	G90654	hypothetical prote	379	29	59.2	1145	2	T18235	probable transcrip
307	29	59.2	452	2	G85505	regulatory protei	380	29	59.2	1199	2	T37561	DNA-directed RNA p
308	29	59.2	452	2	E64745	hypothetical prote	381	29	59.2	1224	2	S73171	hypothetical prote
309	29	59.2	455	2	S50391	regulatory protei	382	29	59.2	1232	2	T47993	hypothetical prote
310	29	59.2	455	2	A70461	hypothetical prote	383	29	59.2	1254	2	T24897	DNA-directed RNA p
311	29	59.2	465	1	F64630	potassium channel	384	29	59.2	1292	2	F64237	glued protein - fr
312	29	59.2	474	2	AC1904	cysteine-tRNA liga	385	29	59.2	1319	2	A28313	probable xanthine
313	29	59.2	475	2	T35799	hypothetical prote	386	29	59.2	1364	2	T51920	myosin-2 isoform -
314	29	59.2	484	2	G64398	probable oxidoredu	387	29	59.2	1471	2	T40117	calcium channel, v
315	29	59.2	484	2	B86155	argininosuccinate	388	29	59.2	1873	2	A55645	vitellogenin I pre
316	29	59.2	491	2	B86155	probable polygalac	389	29	59.2	1912	2	T29088	adherence factor T
317	29	59.2	492	1	S28003	trypanothione-disu	390	29	59.2	3255	2	G81702	hypothetical prote
318	29	59.2	497	2	C98318	alkaline proteinas	391	28	57.1	88	2	T17418	hypothetical prote
319	29	59.2	498	2	F84667	probable cytochrom	392	28	57.1	91	2	T16579	hypothetical prote
320	29	59.2	499	2	T32688	hypothetical prote	393	28	57.1	100	4	S32195	hypothetical prote
321	29	59.2	503	2	AB1933	hypothetical prote	394	28	57.1	112	2	F70403	hypothetical prote
			527	2	T26732	hypothetical prote							

395 28 57.1 113 2 G69065 hypothetical prote
396 28 57.1 115 2 A83554 hypothetical prote
397 28 57.1 122 2 B64493 hypothetical prote
398 28 57.1 124 2 AC0803 probable bacteriop
399 28 57.1 131 2 AG1082 hypothetical prote
400 28 57.1 131 2 AH1439 hypothetical prote
401 28 57.1 134 2 S66294 921-S protein - mo
402 28 57.1 134 2 A57233 complexin I - rat
403 28 57.1 136 2 E70182 hypothetical prote
404 28 57.1 147 2 F70814 hypothetical prote
405 28 57.1 147 2 T26225 hypothetical prote
406 28 57.1 148 2 F75091 hypothetical prote
407 28 57.1 151 2 A84163 hypothetical prote
408 28 57.1 157 2 H75178 hypothetical prote
409 28 57.1 159 2 B84000 hypothetical prote
410 28 57.1 162 2 B81967 probable periplasm
411 28 57.1 162 2 A70489 conserved hypothet
412 28 57.1 165 2 AI0843 conserved hypothet
413 28 57.1 166 2 G81021 thioredoxin, proba
414 28 57.1 172 2 H72742 hypothetical prote
415 28 57.1 173 2 AE1237 conserved hypothet
416 28 57.1 173 2 G64495 hypothetical prote
417 28 57.1 183 2 G70450 hypothetical prote
418 28 57.1 185 2 E95904 hypothetical prote
419 28 57.1 190 2 AH2529 transcription regu
420 28 57.1 191 2 E82715 chorismate mutase
421 28 57.1 196 2 G83541 probable transcrip
422 28 57.1 199 1 S25967 NADH2 dehydrogenas
423 28 57.1 199 1 TVMSB1 transforming prote
424 28 57.1 205 1 TVHUB1 transforming prote
425 28 57.1 207 2 A69349 transforming prote
426 28 57.1 207 2 S56209 conserved hypothet
427 28 57.1 208 1 D71802 probable membrane
428 28 57.1 210 2 A40979 hypothetical prote
429 28 57.1 213 2 A83837 temperature shock-
430 28 57.1 222 2 C97164 siroheme synthase
431 28 57.1 236 1 TVMSA1 uncharacterized pr
432 28 57.1 236 2 JC7383 B-cell lymphoma 2
433 28 57.1 236 2 I67432 BCL-2 - rat (fragm
434 28 57.1 236 2 I53744 gene bcl-2 protein
435 28 57.1 236 2 A82287 phosphate transpor
436 28 57.1 238 2 E97105 probable membrane
437 28 57.1 239 1 TVHUA1 transforming prote
438 28 57.1 243 2 S76271 hypothetical prote
439 28 57.1 246 2 B43912 myogenic regulator
440 28 57.1 255 1 E71090 probable lactam ut
441 28 57.1 255 2 S04105 myogenic factor 5
442 28 57.1 255 2 JN0624 Myogenic factor -
443 28 57.1 255 2 S22825 myf-5 protein - mo
444 28 57.1 256 1 B83852 phosphoesterase-re
445 28 57.1 257 2 E87290 transcription regu
446 28 57.1 258 2 AC2575 chromosome partiti
447 28 57.1 259 2 S41126 Myf5 protein - chi
448 28 57.1 259 2 A86822 hypothetical prote
449 28 57.1 261 2 A95048 enoyl-CoA hydratase
450 28 57.1 261 2 G97918 enoyl-CoA hydratase
451 28 57.1 263 2 D90231 conserved hypothet
452 28 57.1 265 2 AD0620 probable prophage
453 28 57.1 265 2 T40757 conserved hypothet
454 28 57.1 266 2 T35212 probable RNA polym
455 28 57.1 267 2 A69319 thiamin biosynthes
456 28 57.1 268 2 T30630 hypothetical prote
457 28 57.1 272 1 SJGQBG giardin beta chain
458 28 57.1 273 2 E87430 flagellin FljL [im
459 28 57.1 274 2 T16003 hypothetical prote
460 28 57.1 282 2 G84081 hypothetical prote
461 28 57.1 285 2 D70892 hypothetical prote
462 28 57.1 285 2 S51247 thioredoxin homolo
463 28 57.1 291 2 AI1724 ABC transporter (A
464 28 57.1 293 2 AC2480 hypothetical prote
465 28 57.1 297 2 D69327 succinoglycan bios
466 28 57.1 301 1 RGECGT GTP-binding protei
467 28 57.1 301 2 E95977 UTP-glucose-1-phos

468 28 57.1 301 2 E85902 GTP-binding protei
469 28 57.1 301 2 AB0829 GTP-binding protei
470 28 57.1 301 2 AF1257 GTP binding protei
471 28 57.1 301 2 AB1620 GTP-binding protei
472 28 57.1 301 2 H91057 GTP-binding protei
473 28 57.1 302 2 A34406 aldehyde reductase
474 28 57.1 302 2 F64042 GTP-binding protei
475 28 57.1 303 2 E69584 3-methyladenine DN
476 28 57.1 303 2 S45093 Na+/K+-exchanging
477 28 57.1 303 2 A72365 conserved hypothet
478 28 57.1 303 2 AG0331 probable GTP-bindi
479 28 57.1 305 2 F83548 GTP-binding protei
480 28 57.1 305 2 F69306 hypothetical prote
481 28 57.1 307 2 D81935 probable GTPase NM
482 28 57.1 309 2 D88956 protein K04F1.15 [p
483 28 57.1 309 2 H71089 hypothetical prote
484 28 57.1 311 1 IQBS44 primosome componen
485 28 57.1 311 2 C75493 GTP-binding protei
486 28 57.1 312 2 S61213 malate dehydrogena
487 28 57.1 314 2 JT0372 repI protein - Sta
488 28 57.1 316 1 A60603 aldehyde reductase
489 28 57.1 316 2 I49484 aldehyde reductase
490 28 57.1 318 2 C72856 AcOrf-51 protein -
491 28 57.1 319 2 F85054 probable malonyl-C
492 28 57.1 320 2 E91022 o-succinylbenzoyl-
493 28 57.1 320 2 F85866 hypothetical prote
494 28 57.1 320 2 C64997 o-succinylbenzoate
495 28 57.1 323 2 T22492 hypothetical prote
496 28 57.1 324 2 F81169 GTP-binding protei
497 28 57.1 326 2 C64937 hypothetical prote
498 28 57.1 326 2 H90938 hypothetical prote
499 28 57.1 326 2 D85787 hypothetical prote
500 28 57.1 326 2 E98077 UDPglucose 4-epime
501 28 57.1 327 2 A70435 cysteine synthase
502 28 57.1 327 2 D84806 hypothetical prote
503 28 57.1 331 1 A70032 conserved hypothet
504 28 57.1 331 2 T14544 fructokinase (EC 2
505 28 57.1 335 2 H75029 hypothetical prote
506 28 57.1 335 2 AD1314 peptidase homolog
507 28 57.1 335 2 AH3298 sensory transducti
508 28 57.1 336 2 D95213 UDP-glucose 4-epim
509 28 57.1 336 2 T14437 inner envelope mem
510 28 57.1 336 2 H82040 general secretion
511 28 57.1 339 2 E82211 conserved hypothet
512 28 57.1 346 2 F75377 ABC transporter, A
513 28 57.1 351 2 S21345 retrovirus-related
514 28 57.1 357 2 G70010 probable aspartate
515 28 57.1 357 2 A83663 hypothetical prote
516 28 57.1 359 2 T52337 phosphoprotein pho
517 28 57.1 359 2 F84513 hypothetical prote
518 28 57.1 360 2 AG1975 hypothetical prote
519 28 57.1 362 2 D84952 GTP-binding protei
520 28 57.1 370 2 H90559 hypothetical prote
521 28 57.1 370 2 H75318 membrane lipoprote
522 28 57.1 379 2 A57477 potassium channel
523 28 57.1 381 2 T24256 hypothetical prote
524 28 57.1 387 1 D84908 probable phosphoes
525 28 57.1 387 2 T52650 cysteine synthase
526 28 57.1 388 2 F70133 flagellar-associat
527 28 57.1 391 1 LPRTA4 apolipoprotein A-I
528 28 57.1 394 2 B75404 probable amidase -
529 28 57.1 396 2 T11709 nifs protein homol
530 28 57.1 397 2 S46691 exopolyphosphatase
531 28 57.1 403 2 AC1960 two-component resp
532 28 57.1 411 2 AB0199 peptidase T (EC 3.
533 28 57.1 412 2 F97911 probable UDPglucos
534 28 57.1 419 2 E70489 processing protein
535 28 57.1 420 2 A83948 competence-damage
536 28 57.1 426 2 C85755 probable oxidoredu
537 28 57.1 426 2 F90863 probable oxidoredu
538 28 57.1 426 2 H64878 probable oxidoredu
539 28 57.1 430 2 AC2707 two component resp
540 28 57.1 433 2 AI0882 probable oxidoredu

541	28	57.1	433	2	S25194	zuotin - yeast (Sa	614	28	57.1	765	2	T15447	hypothetical prote
542	28	57.1	435	2	AF2582	heat shock chapero	615	28	57.1	777	2	F84985	endopeptidase La (
543	28	57.1	435	2	D97364	hslu protein (AE00	616	28	57.1	783	2	AE1961	Zam protein [impor
544	28	57.1	435	2	F96939	TPR repeats contai	617	28	57.1	786	2	A69308	DNA gyrase, subuni
545	28	57.1	435	2	AI1017	probable type-I se	618	28	57.1	787	2	H86707	formate C-acetyltr
546	28	57.1	435	2	B87592	hypothetical prote	619	28	57.1	787	2	A71642	penicillin-binding
547	28	57.1	435	2	T30114	hypothetical prote	620	28	57.1	788	2	S53397	actin-interacting
548	28	57.1	436	2	C83449	probable oxidoredu	621	28	57.1	804	2	A39972	segment S5 protein
549	28	57.1	437	2	T23652	alpha-1,3-mannosyl	622	28	57.1	808	2	E64492	hypothetical prote
550	28	57.1	437	2	T02689	hypothetical prote	623	28	57.1	832	2	S56230	hypothetical prote
551	28	57.1	439	2	T14872	ABC exporter outer	624	28	57.1	842	2	T16198	hypothetical prote
552	28	57.1	441	1	B33862	transcription regu	625	28	57.1	844	2	C28667	DNA mismatch repai
553	28	57.1	442	2	T47800	cysteine synthase	626	28	57.1	852	2	F44020	hypothetical prote
554	28	57.1	445	2	C97489	response regulator	627	28	57.1	857	2	E98107	DNA mismatch repai
555	28	57.1	449	1	LIPG	triacylglycerol li	628	28	57.1	865	2	A83946	DNA mismatch repai
556	28	57.1	453	2	S52690	hypothetical prote	629	28	57.1	872	2	B89904	DNA mismatch repai
557	28	57.1	455	2	AB0534	membrane-bound lyt	630	28	57.1	874	2	AC3070	ATP-dependent Clp
558	28	57.1	457	2	H85013	hypothetical prote	631	28	57.1	879	2	T19919	hypothetical prote
559	28	57.1	462	2	A64416	adenylosuccinate l	632	28	57.1	884	2	B96492	probable polyprote
560	28	57.1	462	2	T01549	hypothetical prote	633	28	57.1	887	2	F98216	endopeptidase clp
561	28	57.1	465	2	C43357	triacylglycerol li	634	28	57.1	909	2	AC2365	hypothetical prote
562	28	57.1	466	1	P2WLR1	L2 protein - rhesu	635	28	57.1	917	2	S64100	probable membrane
563	28	57.1	467	2	E70976	hypothetical prote	636	28	57.1	919	2	B72263	isoleucine-tRNA li
564	28	57.1	469	1	P2WLJ35	L2 protein - human	637	28	57.1	925	2	T29311	hypothetical prote
565	28	57.1	469	2	S36525	L2 protein - human	638	28	57.1	931	2	T31731	hypothetical prote
566	28	57.1	475	2	AF0132	membrane-bound lyt	639	28	57.1	940	2	D87006	probable preprotei
567	28	57.1	475	2	S20250	splicing factor U2	640	28	57.1	949	2	B70592	probable secA prot
568	28	57.1	480	2	D75050	hypothetical prote	641	28	57.1	955	2	T33040	hypothetical prote
569	28	57.1	484	2	S60943	RUD3 protein - yea	642	28	57.1	956	2	A56920	gliotactin precurs
570	28	57.1	485	2	T03131	alkaline exonuclea	643	28	57.1	968	2	T01733	hypothetical prote
571	28	57.1	492	2	S22646	splicing factor U2	644	28	57.1	968	2	G90636	hypothetical prote
572	28	57.1	495	2	H71933	hypothetical prote	645	28	57.1	968	2	G85487	probable ATP-depen
573	28	57.1	504	2	S73163	H+-transporting tw	646	28	57.1	968	2	C64727	probable ATP-depen
574	28	57.1	505	2	S39520	H+-transporting tw	647	28	57.1	975	1	A31497	kinesin heavy chai
575	28	57.1	509	2	B86667	alkyl hydroperoxid	648	28	57.1	978	1	RGBYI3	regulatory protein
576	28	57.1	509	2	AI2316	competence protein	649	28	57.1	978	2	T40803	probable pre-tRNA
577	28	57.1	509	2	T06300	hypothetical prote	650	28	57.1	1020	2	B82604	cation efflux syst
578	28	57.1	511	2	AG0261	conserved hypothet	651	28	57.1	1044	2	T00342	hypothetical prote
579	28	57.1	515	2	E84577	probable cytokinin	652	28	57.1	1073	2	E83051	carbamoylphosphate
580	28	57.1	517	2	T50801	argininosuccinate	653	28	57.1	1075	2	T34223	hypothetical prote
581	28	57.1	523	1	DEFFG6	glucose-6-phosphat	654	28	57.1	1091	2	F83928	hypothetical prote
582	28	57.1	524	2	A47740	glucose-6-phosphat	655	28	57.1	1155	2	G64332	FUN12/bIF-2 family
583	28	57.1	528	1	A31854	lanosterol 14alpha	656	28	57.1	1221	2	A44978	DNA topoisomerase
584	28	57.1	531	2	T04722	hypothetical prote	657	28	57.1	1252	2	B42771	reticulocyte-bindin
585	28	57.1	540	2	D86432	hypothetical prote	658	28	57.1	1255	2	B97104	hypothetical prote
586	28	57.1	548	1	G69526	methyalmalonyl-CoA	659	28	57.1	1296	2	I40645	botulinum neurotox
587	28	57.1	556	2	C75202	dipeptide abc tran	660	28	57.1	1418	2	T15232	hypothetical prote
588	28	57.1	564	2	T28949	hypothetical prote	661	28	57.1	1436	2	S67655	probable membrane
589	28	57.1	564	2	A87384	hypothetical prote	662	28	57.1	1495	2	S27001	alpha-1-macroglobu
590	28	57.1	574	1	VGNZBS	cell fusion glycop	663	28	57.1	1500	2	A42210	alpha-1-macroglobu
591	28	57.1	574	2	A99189	glycyl-tRNA synthe	664	28	57.1	1553	2	T03301	rab3 effector prot
592	28	57.1	581	2	T29830	hypothetical prote	665	28	57.1	1657	2	T19536	hypothetical prote
593	28	57.1	587	2	A31776	hypothetical prote	666	28	57.1	1787	2	G97222	hypothetical prote
594	28	57.1	592	2	T52139	lactose permease -	667	28	57.1	1846	2	T33079	hypothetical prote
595	28	57.1	602	2	T50974	LRR-containing F-b	668	28	57.1	2150	2	T08165	RNA1 polyprotein -
596	28	57.1	603	1	QOBED1	related to hxB pro	669	28	57.1	2241	2	T20971	hypothetical prote
597	28	57.1	611	2	H82351	HHLF5 protein - hu	670	28	57.1	2261	2	T20978	hypothetical prote
598	28	57.1	619	2	A97156	probable hemolysin	671	28	57.1	2291	1	A46147	spectrin beta chai
599	28	57.1	619	2	F82391	deoxyxylulose-5-ph	672	28	57.1	2685	2	T38755	hypothetical prote
600	28	57.1	626	2	AI3310	ABC transporter, A	673	28	57.1	3068	1	A44062	genome polyprotein
601	28	57.1	639	2	T06735	ABC transporter AT	674	28	57.1	5107	2	T29144	hypothetical prote
602	28	57.1	645	2	F96630	hypothetical prote	675	28	57.1	5107	2	T21989	hypothetical prote
603	28	57.1	653	2	E84206	hypothetical prote	676	27.5	56.1	433	2	A83870	hypothetical prote
604	28	57.1	657	2	G71602	methyalmalonyl-CoA	677	27	55.1	36	2	A83870	complement C5a - r
605	28	57.1	657	2	G71602	protein with DnaJ	677	27	55.1	77	2	A57689	hypothetical prote
606	28	57.1	668	2	C71496	probable exinuclea	678	27	55.1	78	2	B75193	hypothetical prote
607	28	57.1	669	2	C96814	hypothetical prote	679	27	55.1	78	2	B90999	hypothetical prote
608	28	57.1	676	2	D81654	excinuclease ABC c	680	27	55.1	93	2	G91172	probable phosphotr
609	28	57.1	689	2	E89898	hypothetical prote	681	27	55.1	93	2	G86018	probable phosphotr
610	28	57.1	701	1	B44259	kinesin-related pr	682	27	55.1	95	2	AG0964	probable PTS syste
611	28	57.1	701	2	T20892	hypothetical prote	683	27	55.1	101	2	H72542	hypothetical prote
612	28	57.1	707	2	T29559	hypothetical prote	684	27	55.1	102	2	E97788	hypothetical prote
613	28	57.1	754	2	H84710	Mutator-like trans	685	27	55.1	105	1	CCMST	cytochrome c, test
	28	57.1	760	2	T28224	ORF MSV063 probabl	686	27	55.1	107	2	T42918	hypothetical prote

100

687 27 55.1 107 2 T42903 hypothetical prote
688 27 55.1 108 2 S61627 hypothetical prote
689 27 55.1 108 2 T24634 hypothetical prote
690 27 55.1 112 2 D90078 conserved hypothet
691 27 55.1 113 2 C71223 hypothetical prote
692 27 55.1 115 2 C64578 conserved hypothet
693 27 55.1 118 2 B64630 hypothetical prote
694 27 55.1 121 2 C96932 transcription regu
695 27 55.1 123 2 E81851 ribosomal-binding f
696 27 55.1 123 2 B81091 ribosome-binding f
697 27 55.1 124 2 C90423 hypothetical prote
698 27 55.1 126 2 E86778 regulatory protein
699 27 55.1 129 2 AB1725 transcription regu
700 27 55.1 131 2 AI1354 transcription regu
701 27 55.1 133 2 A75314 ribosomal protein
702 27 55.1 133 2 F90037 hypothetical prote
703 27 55.1 133 2 AC0427 conserved hypothet
704 27 55.1 134 1 JQ1889 AL3 protein - toma
705 27 55.1 135 2 G64395 hypothetical prote
706 27 55.1 137 2 G95102 hypothetical prote
707 27 55.1 138 1 C64005 hypothetical prote
708 27 55.1 140 2 A48552 protein US420 - in
709 27 55.1 151 2 H69078 ferripyochelin bin
710 27 55.1 151 2 T01176 hypothetical prote
711 27 55.1 154 2 T17580 hypothetical prote
712 27 55.1 157 2 AE2434 hypothetical prote
713 27 55.1 158 1 S38990 glycine reductase
714 27 55.1 158 2 G70031 mutator MutT prote
715 27 55.1 166 2 C71874 hypothetical prote
716 27 55.1 175 2 E89992 hypothetical prote
717 27 55.1 175 2 AB1657 hypothetical prote
718 27 55.1 175 2 AG1285 hypothetical prote
719 27 55.1 177 2 T16280 hypothetical prote
720 27 55.1 177 2 AG1831 hypothetical prote
721 27 55.1 178 2 H70345 conserved hypothet
722 27 55.1 181 2 A70178 conserved hypothet
723 27 55.1 182 2 I67436 interleukin-1-beta
724 27 55.1 183 1 NKVLCP core antigen - hep
725 27 55.1 185 2 C87646 transcription anti
726 27 55.1 186 1 Q3ECTR hypothetical 20.6K
727 27 55.1 186 2 H91128 hypothetical prote
728 27 55.1 186 2 G85973 hypothetical prote
729 27 55.1 186 2 B69389 conserved hypothet
730 27 55.1 188 2 B82584 transcription regu
731 27 55.1 188 2 AI1225 hypothetical prote
732 27 55.1 188 2 F82956 conserved hypothet
733 27 55.1 189 1 IVBOI1 interferon alpha-I
734 27 55.1 189 1 IVBOIC interferon alpha-I
735 27 55.1 189 1 IVBOID interferon alpha-I
736 27 55.1 190 2 E90143 conserved hypothet
737 27 55.1 192 2 S12314 hypothetical prote
738 27 55.1 193 2 D70046 conserved hypothet
739 27 55.1 193 2 S10911 hypothetical prote
740 27 55.1 199 2 T19584 hypothetical prote
741 27 55.1 200 2 T23485 hypothetical prote
742 27 55.1 203 2 D86660 30S ribosomal prot
743 27 55.1 204 2 T10766 patatin-like latex
744 27 55.1 206 2 AH1588 hypothetical prote
745 27 55.1 206 2 AC1651 hypothetical prote
746 27 55.1 206 2 AC3478 protein ybIs precu
747 27 55.1 207 2 T37464 probable glutathio
748 27 55.1 212 2 T31759 hypothetical prote
749 27 55.1 214 1 KIECA adenylate kinase (
750 27 55.1 214 2 AC0563 adenylate kinase (
751 27 55.1 214 2 G90694 adenylate kinase (
752 27 55.1 214 2 C85545 adenylate kinase (
753 27 55.1 219 2 C70659 probable lipoprote
754 27 55.1 220 2 D70659 probable lipoprote
755 27 55.1 221 2 B70378 transcription regu
756 27 55.1 221 2 C64040 hypothetical prote
757 27 55.1 224 2 F69601 cytidylate kinase
758 27 55.1 226 2 AF2362 transcription regu
759 27 55.1 227 2 F72334 conserved hypothet

760 27 55.1 234 2 E86622 hypothetical prote
761 27 55.1 234 2 E72001 hypothetical prote
762 27 55.1 234 2 AI3170 IS5 family transpo
763 27 55.1 236 2 F97736 hypothetical prote
764 27 55.1 236 2 D71733 phosphoribosylamin
765 27 55.1 240 1 A70031 conserved hypothet
766 27 55.1 240 2 H75014 3-oxoacyl-[acyl-ca
767 27 55.1 241 2 A75065 hypothetical prote
768 27 55.1 242 2 A97117 tRNA-(guanine-N1)-
769 27 55.1 243 2 F81536 hypothetical prote
770 27 55.1 244 2 H69844 diadenosine tetrap
771 27 55.1 244 2 S73750 ATP-binding protei
772 27 55.1 252 2 AC1239 undecaprenyl diph
773 27 55.1 252 2 AG1601 undecaprenyl diph
774 27 55.1 253 2 D75193 hypothetical prote
775 27 55.1 254 1 DFHU beta-neoendorphin
776 27 55.1 255 1 SNRTC8 proteasome endopep
777 27 55.1 256 1 DFPG beta-neoendorphin
778 27 55.1 257 2 H70198 thymidylate kinase
779 27 55.1 258 2 AF2287 group 3 sigma 37-t
780 27 55.1 258 2 JC6318 beta-neoendorphin
781 27 55.1 259 1 CTXLPA corticotropin / li
782 27 55.1 260 1 CTXLPB corticotropin / li
783 27 55.1 261 2 B42882 motility protein (
784 27 55.1 264 2 G70195 pyridoxal kinase (
785 27 55.1 266 1 F64414 indole-3-glycerol-
786 27 55.1 266 2 AH3017 extragenic suppres
787 27 55.1 266 2 H98266 shikimate 5-dehydr
788 27 55.1 269 2 D90173 hypothetical prote
789 27 55.1 269 2 T18145 UTP-glucose-1-phos
790 27 55.1 273 2 E71913 flagellin FljO [im
791 27 55.1 273 2 G87347 hypothetical prote
792 27 55.1 275 2 T48436 purine nucleoside
793 27 55.1 277 2 C70351 cell division inhi
794 27 55.1 278 2 H70404 ABC-type sulfate t
795 27 55.1 279 2 A96913 flagellar synthesi
796 27 55.1 280 2 B83463 conserved hypothet
797 27 55.1 281 2 AH1738 acetyl-coenzyme A
798 27 55.1 289 2 E71875 acetyl-CoA carboxy
799 27 55.1 289 2 F64638 conserved hypothet
800 27 55.1 291 2 AB1384 conserved hypothet
801 27 55.1 291 2 AD1759 conserved hypothet
802 27 55.1 291 2 AD1303 conserved hypothet
803 27 55.1 291 2 AD1675 phage-related prot
804 27 55.1 292 2 G84092 adrenocortical est
805 27 55.1 296 2 A44011 homoserine kinase
806 27 55.1 297 2 A70413 hypothetical prote
807 27 55.1 301 2 E96630 hypothetical prote
808 27 55.1 301 2 T51435 hypothetical prote
809 27 55.1 301 2 A70335 hypothetical prote
810 27 55.1 308 2 E84374 prephenate dehydro
811 27 55.1 309 2 AI3000 tyrC protein [impo
812 27 55.1 310 2 G98282 hypothetical prote
813 27 55.1 311 2 H98278 hypothetical prote
814 27 55.1 311 2 AI3004 hypothetical prote
815 27 55.1 312 2 I40383 malate dehydrogena
816 27 55.1 313 2 C96926 endoglucanase, ami
817 27 55.1 313 2 H64144 hypothetical prote
818 27 55.1 314 2 S00937 repE protein - Sta
819 27 55.1 314 2 AG3108 conserved hypothet
820 27 55.1 316 2 T08807 mitochondrial proc
821 27 55.1 317 2 T25560 hypothetical prote
822 27 55.1 318 2 B72676 hypothetical prote
823 27 55.1 319 2 T41795 AcMNPV orf51 - Bom
824 27 55.1 319 2 A47499 thiamine diphospho
825 27 55.1 321 2 F82296 conserved hypothet
826 27 55.1 324 1 S73340 hypothetical prote
827 27 55.1 324 2 AE1364 protein gp49 [Bact
828 27 55.1 326 2 E98178 hypothetical prote
829 27 55.1 327 2 G90139 deacetylase, proba
830 27 55.1 327 2 T34562 G protein pathway
831 27 55.1 330 2 A70422 hypothetical prote
832 27 55.1 331 2 A69113 isocitrate dehydro

833	27	55.1	331	2	A12972	two component sens	906	27	55.1	438	2	E55578	hypothetical prote
834	27	55.1	331	2	B98310	probable transmemb	907	27	55.1	441	2	F86258	protein F5011.14 [
835	27	55.1	332	2	C64073	rbs repressor homo	908	27	55.1	444	2	B81825	glutamate dehydrog
836	27	55.1	332	2	T28882	SUP-10 precursor -	909	27	55.1	444	2	H81050	glutamate dehydrog
837	27	55.1	337	2	C90400	hypothetical prote	910	27	55.1	444	2	B69530	adenylosuccinate 1
838	27	55.1	338	2	H72679	hypothetical prote	911	27	55.1	446	2	D82730	adenosylhomocystei
839	27	55.1	338	2	T23365	hypothetical prote	912	27	55.1	446	2	A83213	probable ATP-depen
840	27	55.1	341	2	D64159	hypothetical prote	913	27	55.1	448	2	G89933	hypothetical prote
841	27	55.1	342	2	D95906	probable transcrip	914	27	55.1	449	2	G81419	hypothetical prote
842	27	55.1	342	2	T43410	replication factor	915	27	55.1	453	1	FGHUGB	fibrinogen gamma-B
843	27	55.1	342	2	S42885	beta-1,3-glucanase	916	27	55.1	458	2	G83690	hypothetical prote
844	27	55.1	342	2	S23438	hypothetical prote	917	27	55.1	460	2	F83953	hypothetical prote
845	27	55.1	342	2	T23224	hypothetical prote	918	27	55.1	465	2	C75028	h+-transporting AT
846	27	55.1	345	2	E70484	recombination prot	919	27	55.1	465	2	F71213	probable H(+)-tran
847	27	55.1	352	2	S38147	hypothetical prote	920	27	55.1	466	2	E70183	purine-binding che
848	27	55.1	352	2	D72264	hypothetical prote	921	27	55.1	467	1	P2WL33	L2 protein - human
849	27	55.1	352	2	C98155	probable DNA-bindi	922	27	55.1	467	2	B72331	hypothetical prote
850	27	55.1	353	2	E70100	cryptophan-tRNA li	923	27	55.1	469	2	AE0835	probable type I se
851	27	55.1	353	2	A81737	UDP-N-acetylglucos	924	27	55.1	473	1	P2WLHS	minor capsid prote
852	27	55.1	357	2	A13132	transcription regu	925	27	55.1	480	2	C69853	probable tagaturon
853	27	55.1	360	2	E69086	cell division prot	926	27	55.1	483	2	A85723	altronate oxidored
854	27	55.1	365	2	B54128	Fc-binding protein	927	27	55.1	483	2	D64906	tagaturonate reduc
855	27	55.1	366	2	G71179	probable acetylorn	928	27	55.1	483	2	H90894	altronate oxidored
856	27	55.1	367	2	T06131	isocitrate dehydro	929	27	55.1	484	2	AD0699	probable pathogeni
857	27	55.1	370	2	D83793	hypothetical prote	930	27	55.1	486	2	T37572	probable metallopr
858	27	55.1	371	2	B85090	hypothetical prote	931	27	55.1	491	2	T23524	hypothetical prote
859	27	55.1	375	2	I39169	sialyltransferase	932	27	55.1	492	2	F70326	conserved hypothet
860	27	55.1	375	2	I48686	N-glycan alpha 2,8	933	27	55.1	493	2	G84594	probable diacylgly
861	27	55.1	375	2	A46727	sialyltransferase	934	27	55.1	494	2	G64382	acetolactate synth
862	27	55.1	375	2	S13025	NADH2 dehydrogenas	935	27	55.1	494	2	T28660	probable adhesin P
863	27	55.1	378	2	T00481	probable RING zinc	936	27	55.1	499	2	E84776	hypothetical prote
864	27	55.1	382	1	WZVZ16	I6 43.5K protein -	937	27	55.1	503	2	S78321	H+-transporting tw
865	27	55.1	382	2	B72158	L6L protein - vari	938	27	55.1	506	2	S71591	aspartic proteinas
866	27	55.1	382	2	S33075	I6L protein - vari	939	27	55.1	511	2	S58399	cellutagmin I sytV
867	27	55.1	382	2	T28498	hypothetical prote	940	27	55.1	513	2	G83725	GMP synthetase gua
868	27	55.1	383	2	T50502	hypothetical prote	941	27	55.1	520	1	A41771	3-oxoacid CoA-tran
869	27	55.1	383	2	E81156	histidyl-tRNA synt	942	27	55.1	521	2	D64434	serine-tRNA ligase
870	27	55.1	383	2	E81950	hypothetical prote	943	27	55.1	521	2	O4CK51	lanosterol 14alpha
871	27	55.1	386	2	C90190	tryptophanyl-tRNA	944	27	55.1	528	1	G72605	probable acid-CoA
872	27	55.1	392	2	S64914	hypothetical prote	945	27	55.1	529	2	H71450	probable DNA repai
873	27	55.1	397	2	A72408	conserved hypothet	946	27	55.1	529	2	D71338	probable ribose/ga
874	27	55.1	398	2	G71733	hypothetical prote	947	27	55.1	533	1	E86412	hypothetical prote
875	27	55.1	399	2	T11954	carbamoyl phosphat	948	27	55.1	533	2	T06683	aldehyde dehydroge
876	27	55.1	400	2	D72277	phosphoribosylamin	949	27	55.1	539	2	A99481	hypothetical prote
877	27	55.1	401	2	T19258	hypothetical prote	950	27	55.1	540	2	T21647	hypothetical prote
878	27	55.1	402	2	C95035	conserved hypothet	951	27	55.1	541	2	S51799	nucleoporin NUP57
879	27	55.1	403	2	T39846	probable nuclear p	952	27	55.1	546	2	A99600	conserved hypothet
880	27	55.1	404	2	F82092	aminotransferase,	953	27	55.1	548	2	E81698	pyrophosphate-fruc
881	27	55.1	405	2	E90154	primase (Amino-end	954	27	55.1	548	2	G82286	phosphate ABC tran
882	27	55.1	406	2	F70019	nifs protein homol	955	27	55.1	549	2	F64640	conserved hypothet
883	27	55.1	411	2	I39809	creatinase (EC 3.5	956	27	55.1	553	2	B64622	osmoprotection pro
884	27	55.1	413	2	F70211	hypothetical prote	957	27	55.1	555	2	D95377	probable cerebrosi
885	27	55.1	416	2	A48249	pre-mRNA splicing	958	27	55.1	556	2	T23172	hypothetical prote
886	27	55.1	419	2	A27539	variant surface gl	959	27	55.1	557	2	T45135	chaperone protein
887	27	55.1	420	2	S76601	hypothetical prote	960	27	55.1	565	2	T29718	hypothetical prote
888	27	55.1	420	2	AF2464	hypothetical prote	961	27	55.1	565	2	JE0160	sodium bicarbonate
889	27	55.1	420	2	C69184	hypothetical prote	962	27	55.1	567	1	D43719	urease (EC 3.5.1.5
890	27	55.1	421	2	H89916	hypothetical prote	963	27	55.1	567	1	S08480	urease (EC 3.5.1.5
891	27	55.1	421	2	AE2473	diaminopimelate de	964	27	55.1	570	2	T38148	phosphatidyl synth
892	27	55.1	423	2	D81248	hypothetical prote	965	27	55.1	572	2	S68119	laccase (EC 1.10.3
893	27	55.1	423	2	AC3553	3-deoxy-D-manno-oc	966	27	55.1	575	2	S44018	aerobactin biosynt
894	27	55.1	425	2	T46698	4-aminobutyrate tr	967	27	55.1	576	2	S68117	laccase (EC 1.10.3
895	27	55.1	427	1	A39744	hypothetical prote	968	27	55.1	576	2	T38666	probable trp-asp r
896	27	55.1	430	2	QJ1392	citrate (si)-synth	969	27	55.1	581	2	H70125	hypothetical prote
897	27	55.1	430	2	T37198	hypothetical prote	970	27	55.1	582	2	B70389	translation elonga
898	27	55.1	431	2	JC7730	1-aminocyclopropan	971	27	55.1	582	2	D82146	transport ATP-bind
899	27	55.1	432	2	T33118	hypothetical prote	972	27	55.1	583	1	H71326	probable protein-e
900	27	55.1	434	2	F69177	hypothetical prote	973	27	55.1	584	2	B72362	acetolactate synth
901	27	55.1	434	2	T43197	nucleoporin homolo	974	27	55.1	588	2	A82232	malate oxidoreduct
902	27	55.1	435	2	B72418	conserved hypothet	975	27	55.1	592	2	T34446	hypothetical prote
903	27	55.1	437	1	FGHUG	fibrinogen gamma-A	976	27	55.1	596	2	F82174	methyl-accepting c
904	27	55.1	437	2	AH1271	hypothetical prote	977	27	55.1	598	2	T01720	hypothetical prote
905	27	55.1	437	2	AB1634	hypothetical prote	978	27	55.1	600	2	C97790	hypothetical prote

979 27 55.1 601 2 AFL192 heat shock protein
980 27 55.1 611 1 S62811 oligoendopeptidase
981 27 55.1 611 2 E90543 glucose inhibited
982 27 55.1 622 2 F97054 Fe-S oxidoreductas
983 27 55.1 623 2 T07933 polyadenylate-bind
984 27 55.1 632 2 B69310 mRNA 3'-end proces
985 27 55.1 632 2 S58152 hypothetical prote
986 27 55.1 639 2 H97807 excinuclease ABC s
987 27 55.1 640 2 T29784 hypothetical prote
988 27 55.1 644 2 S51765 sulfate transport
989 27 55.1 647 2 F95189 threonyl-trRNA synt
990 27 55.1 657 2 AI1882 nitrate transport
991 27 55.1 658 2 A33598 endoglucanase 3 (E
992 27 55.1 659 2 C85016 hypothetical prote
993 27 55.1 660 2 G98055 threonine-trRNA lig
994 27 55.1 663 2 F82160 conserved hypothet
995 27 55.1 663 2 A33618 heat shock protein
996 27 55.1 669 2 D72278 endo-1,4-beta-mann
997 27 55.1 673 2 T06294 hypothetical prote
998 27 55.1 677 2 G69895 formate dehydrogen
999 27 55.1 682 2 B84021 two-component sens
1000 27 55.1 686 2 D71292 probable DNA recom

ALIGNMENTS

RESULT 1
G71827
hypothetical protein jhp1271 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: G71827
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: G71827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1164 <ARN>
A;Cross-references: UNIPROT:Q9ZJN3; UNIPARC:UPI00000D72F4; GB:AE001550; GB:AE001439; NID
A;Experimental source: strain J99

Query Match 81.6%; Score 40; DB 2; Length 1164;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:||||:|
Db 740 ESLKEQVLKY 749

RESULT 2
S73600
probable GTP-binding protein spg - Mycoplasma pneumoniae (strain ATCC 29342)
N;Alternate names: hypothetical protein D02_orf291
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 31-Dec-2004
C;Accession: S73600
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73600
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-291 <HIM>
A;Cross-references: UNIPROT:P75210; UNIPARC:UPI000012A0F3; EMBL:AE000027; GB:U00089; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C;Genetics:
A;Gene: spg
A;Genetic code: GSGC3
C;Superfamily: GTP-binding protein era homolog, bacteria type; translation elongation fa
C;Keywords: nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)

Query Match 77.6%; Score 38; DB 2; Length 291;
Best Local Similarity 60.0%; Pred. No. 9.2;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:||||:|
Db 191 EALREQIIKY 200

RESULT 3
A47305
translation initiation factor eIF-5 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47305
R;Das, K.; Chevesich, J.; Maitra, U.
Proc. Natl. Acad. Sci. U.S.A. 90, 3058-3062, 1993
A;Title: Molecular cloning and expression of cDNA for mammalian translation initiation f
A;Reference number: A47305; MUID:93219424; PMID:8464924
A;Accession: A47305
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-429 <DAS>
A;Cross-references: UNIPROT:Q07205; UNIPARC:UPI000012D3A8; GB:L11651; NID:G294544; PIDN:
A;Experimental source: insulinoma
A;Note: sequence extracted from NCBI backbone (NCBIN:128800, NCBIP:128802)

Query Match 75.5%; Score 37; DB 2; Length 429;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:||||:|
Db 285 EKIREQIKKY 294

RESULT 4
S54048
ribosomal protein S13.e, cytosolic - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D4252; protein YBRC151i; protein YD9609.18; protein YDR064w;
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S54048; S58844; S61753; S11578; S67880
R;Hunt, S.; Bowman, S.; Harris, D.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54031
A;Accession: S54048
A;Molecule type: DNA
A;Residues: 1-151 <HUN>
A;Cross-references: UNIPROT:P05756; UNIPARC:UPI000004F8EC; EMBL:Z49209; NID:G798897; PID:
R;Brandt, P.; Otto, B.; Ramlow, S.; Blocker, H.
submitted to the EMBL Data Library, January 1995
A;Reference number: S58832
A;Accession: S58844
A;Molecule type: DNA
A;Residues: 1-151 <BRA>
A;Cross-references: UNIPARC:UPI000004F8EC; EMBL:X84162; NID:G706817; PIDN:CAA58980.1; PII
R;Brandt, P.; Ramlow, S.; Otto, B.; Bloecker, H.
Yeast 12, 85-90, 1996
A;Title: Nucleotide sequence analysis of a 32,500 bp region of the right arm of Saccharo
A;Reference number: S61741; MUID:96381250; PMID:8789263
A;Accession: S61753
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <BRW>
A;Cross-references: UNIPARC:UPI000004F8EC; EMBL:X84162; NID:G706817; PIDN:CAA58980.1; PII

R;Otaka, E.; Higo, K.; Osawa, S.
Biochemistry 21, 4545-4550, 1982
A;Title: Isolation of seventeen proteins and amino-terminal amino acid sequences of eight
A;Reference number: S11575; MUID:83048950; PMID:6814480
A;Accession: S11578
A;Molecule type: protein
A;Residues: 2-24,'G',26-31,'C',33-35,'X',37-41 <OTA>
A;Cross-references: UNIPARC:UPI00001772A5
R;Bloecker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67587
A;Accession: S67880
A;Molecule type: DNA
A;Residues: 1-151 <BLO>
A;Cross-references: UNIPARC:UPI000004F8EC; EMBL:Z74360; MIPS:YDR064w; NID:gl431516; PIDN:
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:RPS13C
A;Cross-references: SGD:S0002471; MIPS:YDR064w
A;Map position: 4R
A;Introns: 7/3
C;Superfamily: rat ribosomal protein S13; eubacterial ribosomal protein S15 homology
C;Keywords: protein biosynthesis; ribosome
F;82-148/Domain: eubacterial ribosomal protein S15 homology <ES15>

Query Match 73.5%; Score 36; DB 2; Length 151;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 31 ESVIEQIVKY 40

RESULT 5
F64816
ybiA protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: F64816
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F64816
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-160 <BLAT>
A;Cross-references: UNIPROT:P30176; UNIPARC:UPI000013A364; GB:AE000182; GB:U00096; NID:9
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybiA

Query Match 73.5%; Score 36; DB 2; Length 160;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
|||:||||:
Db 90 ESVKEQVMR 98

RESULT 6
T30254
jumonji protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30254
R;Takeuchi, T.; Yamazaki, Y.; Katoh-Fukui, Y.; Tsuchiya, R.; Kondo, S.; Motoyama, J.; Higuchi
Genes Dev. 9, 1211-1222, 1995
A;Title: Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation
A;Reference number: Z20792; MUID:95278734; PMID:7758946

A;Accession: T30254
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1234 <TAK>
A;Cross-references: UNIPROT:Q62315; UNIPARC:UPI000000406B; EMBL:D31967; NID:g780143; PID:
A;Experimental source: strain 129/Ola
C;Genetics:
A;Gene: jumonji
C;Function:
A;Description: may be essential for normal morphogenesis of the neural tube

Query Match 73.5%; Score 36; DB 2; Length 1234;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 573 ESVRAQVEKY 582

RESULT 7
A29924
acetyl-CoA carboxylase (EC 6.4.1.2), hepatic - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A29924; A29337; A27903
R;Takai, T.; Yokoyama, C.; Wada, K.; Tanabe, T.
J. Biol. Chem. 263, 2651-2657, 1988
A;Title: Primary structure of chicken liver acetyl-CoA carboxylase deduced from cDNA sequence
A;Reference number: A29924; MUID:88139305; PMID:2893793
A;Accession: A29924
A;Molecule type: mRNA
A;Residues: 1-2324 <TAK1>
A;Cross-references: UNIPROT:P11029; UNIPARC:UPI0000127CAB; GB:J03541; NID:g211567; PIDN:
R;Takai, T.; Wada, K.; Tanabe, T.
FEBS Lett. 212, 98-102, 1987
A;Title: Primary structure of the biotin-binding site of chicken liver acetyl-CoA carboxylase
A;Reference number: A91375; MUID:87106011; PMID:2879745
A;Accession: A29337
A;Molecule type: mRNA
A;Residues: 493-820 <TAK2>
A;Cross-references: UNIPARC:UPI0000171282; GB:X05019; NID:g63021; PIDN:CAA28675.1; PID:9
A;Accession: A27903
A;Molecule type: mRNA
A;Residues: 493-552,554-783,'RSPS',789-820 <TAK3>
A;Cross-references: UNIPARC:UPI000017312E
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase; liver
F;120-620/Domain: biotin carboxylase homology <BCH>
F;747-819/Domain: lipoyl/biotin-binding homology <LPB>
F;786/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 1; Length 2324;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 1467 ESVRSMVMRY 1476

RESULT 8
A35578
acetyl-CoA carboxylase (EC 6.4.1.2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A35578; A37119; I59145; I70069; I70070; I55305
R;Lopez-Casillas, F.; Bai, D.H.; Luo, X.; Kong, I.S.; Hermodson, M.A.; Kim, K.H.
Proc. Natl. Acad. Sci. U.S.A. 85, 5784-5788, 1988
A;Title: Structure of the coding sequence and primary amino acid sequence of acetyl-coe
A;Reference number: A35578; MUID:88320328; PMID:2901088
A;Accession: A35578
A;Molecule type: mRNA

A;Residues: 1-2345 <LOP>
A;Cross-references: UNIPROT:P11497; UNIPARC:UPI0000127C7B; GB:J03808; NID:g202644; PIDN:
R;Kong, I.S.; Lopez-Casillas, F.; Kim, K.H.
J. Biol. Chem. 265, 13695-13701, 1990
A;Title: Acetyl-CoA carboxylase mRNA species with or without inhibitory coding sequence
A;Reference number: A37119; MUID:90337981; PMID:1974251
A;Accession: A37119
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1167-1200 <KON>
A;Cross-references: UNIPARC:UPI000017312F; GB:M55315
A;Experimental source: clone lambdaDHN121
R;Luo, X.
Proc. Natl. Acad. Sci. U.S.A. 86, 4042-4046, 1989
A;Title: Structural features of the acetyl-CoA carboxylase gene: Mechanism for the gene
A;Reference number: I59145; MUID:89264558; PMID:2566999
A;Accession: I59145
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-74 <RES>
A;Cross-references: UNIPARC:UPI0000170BB7; GB:M26731; NID:g202641; PIDN:AAA40652.1; PID:
A;Experimental source: hepatic
R;Lopez-Casillas, F.; Kim, K.
J. Biol. Chem. 264, 7176-7184, 1989
A;Title: Heterogeneity at the 5' end of rat acetyl-coenzyme A carboxylase mRNA: Lipogeni
A;Reference number: I55305; MUID:89214151; PMID:2565337
A;Accession: I70069
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE2>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26196; NID:g202649; PIDN:AAA40655.1; PID:
A;Accession: I70070
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE3>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26197; NID:g202651; PIDN:AAA40656.1; PID:
A;Accession: I55305
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE4>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26195; NID:g202647; PIDN:AAA40654.1; PID:
A;Experimental source: hepatic
C;Comment: This enzyme catalyzes the carboxylation of acetyl CoA to malonyl CoA and is b
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase
F;119-619/Domain: biotin carboxylase homology <BCH>
F;746-818/Domain: lipoyl/biotin-binding homology <LPB>
F;785/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 1; Length 2345;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||
Db 1489 ESVRSMVMRY 1498

RESULT 9
I38928
acetyl-CoA carboxylase (EC 6.4.1.2), hepatic - human
N;Alternate names: acetyl-Coenzyme A carboxylase
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: I38928
R;Abu-Elheiga, L.; Jayakumar, A.; Baldini, A.; Chirala, S.S.; Wakil, S.J.
Proc. Natl. Acad. Sci. U.S.A. 92, 4011-4015, 1995
A;Title: Human acetyl-CoA carboxylase: characterization, molecular cloning, and evidence
A;Reference number: I38928; MUID:95249602; PMID:7732023
A;Accession: I38928
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2346 <RES>

A;Cross-references: UNIPROT:Q13085; UNIPARC:UPI0000127C6D; EMBL:U19822; NID:g849082; PID
A;Experimental source: HepG2 cells
C;Genetics:
A;Gene: ACC
A;Map position: 17q12
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase; liver
F;120-620/Domain: biotin carboxylase homology <BCH>
F;747-819/Domain: lipoyl/biotin-binding homology <LPB>
F;786/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||
Db 1490 ESVRSMVMRY 1499

RESULT 10
G72358
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: G72358
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: G72358
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-486 <ARN>
A;Cross-references: UNIPROT:Q9WZ51; UNIPARC:UPI00000C13E8; GB:AE001733; GB:AE000512; NID
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0582
C;Superfamily: hypothetical protein MJ0966

Query Match 71.4%; Score 35; DB 2; Length 486;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||
Db 439 ESRESVVQY 448

RESULT 11
S45907
myb-related protein REB1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YBR049c; protein YBR0502
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: S45907; S49504; A35924; S55847
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S45907
A;Molecule type: DNA
A;Residues: 1-810 <ALJ>
A;Cross-references: UNIPROT:P21538; UNIPARC:UPI0000053008; EMBL:Z35918; NID:g536279; PID
A;Experimental source: strain S288C
R;Aljinovic, G.
submitted to the EMBL Data Library, October 1994
A;Description: Sequence and analysis of 24 kb on chromosome II of Saccharomyces cerevisiae
A;Reference number: S49503
A;Accession: S49504
A;Molecule type: DNA
A;Residues: 1-810 <AL2>

A;Cross-references: UNIPARC:UPI0000053008; EMBL:Z46260; NID:g559942; PIDN:CAA86391.1; PI
A;Experimental source: strain S288C
R;Ju, Q.; Morrow, B.E.; Warner, J.R.
Mol. Cell. Biol. 10, 5226-5234, 1990
A;Title: REB1, a yeast DNA-binding protein with many targets, is essential for cell grow
A;Reference number: A35924; MUID:90377212; PMID:2204808
A;Accession: A35924
A;Molecule type: DNA
A;Residues: 1-55, 'N', 57-591, 'RAVVFKNNNNFFNKSSKMMTMLRSC', 619-626, 'E', 628-635, 'E', 637-810
A;Cross-references: UNIPARC:UPI0000168DD6; EMBL:M58728; NID:gl72371; PIDN:AAA34963.1; PI
R;Aljinovic, G.; Pohl, T.M.
Yeast 11, 475-479, 1995
A;Title: Sequence and analysis of 24 kb on chromosome II of Saccharomyces cerevisiae.
A;Reference number: S55846; MUID:95321020; PMID:7597852
A;Accession: S55847
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-810 <ALW>
A;Cross-references: UNIPARC:UPI0000053008; EMBL:Z46260; NID:g559942; PIDN:CAA86391.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C;Genetics:
A;Gene: SGD:REB1; MIPS:YBR049C
A;Cross-references: SGD:S0000253; MIPS:YBR049C
A;Map position: 2R
C;Superfamily: myb-related protein REB1; myb DNA-binding repeat homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;470-519/Domain: myb DNA-binding repeat homology <MYB>

Query Match 71.4%; Score 35; DB 1; Length 810;
Best Local Similarity 70.0%; Pred. No. 98;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| | | | |
Db 13 ESVEEAVLKY 22

RESULT 12
E69681
peptide synthetase ppsE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E69681
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69681
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1279 <KUN>
A;Cross-references: UNIPROT:O31827; UNIPARC:UPI00000604A5; GB:Z99113; GB:AL009126; NID:9
A;Experimental source: strain 168
C;Genetics:
A;Gene: ppsE
C;Superfamily: peptide synthetase ppsE; acetate-CoA ligase homology; acyl carrier protei
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;522-960/Domain: acetate-CoA ligase homology <ACL>
F;978-1046/Domain: acyl carrier protein homology <ACP>
F;1010/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 71.4%; Score 35; DB 1; Length 1279;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
|:|:|:|:|
Db 1180 EAVREKVMK 1188

RESULT 13
G82176
probable outer membrane protein VC1622 [imported] - Vibrio cholerae (strain N16961 serog
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82176
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82176
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <HEI>
A;Cross-references: UNIPROT:Q9KRL5; UNIPARC:UPI00000C3053; GB:AE004240; GB:AE003852; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1622
A;Map position: 1

Query Match 69.4%; Score 34; DB 2; Length 212;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:|
Db 147 EAVRDQLRY 156

RESULT 14
B90544
hypothetical protein MYPV 2580 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 05-Oct-2004
C;Accession: B90544
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.,
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: B90544
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-272 <KUR>
A;Cross-references: UNIPROT:Q98QV3; UNIPARC:UPI00000D45A0; GB:AL4455566; PID:gl4089671; F
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPV 2580
A;Genetic code: SGC3
C;Superfamily: bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrof

Query Match 69.4%; Score 34; DB 2; Length 272;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|:|:|
Db 236 ESIKEALKY 245

RESULT 15
AD1062
protein kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain C

C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD1062
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD1062
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <PAR>
A;Cross-references: UNIPARC:UPI000005A993; GB:AL513382; PIDN:CAD06945.1; PID:gl6505592;
C;Genetics:
A;Gene: STY4823

Query Match 69.4%; Score 34; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SVREQVMKY 10
|||:|
Db 254 SVRDQIAKY 262

RESULT 16
E90514
hypothetical protein MYPV_0210 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: E90514
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: E90514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-455 <KUR>
A;Cross-references: UNIPROT:Q98RI8; UNIPARC:UPI00000D4547; GB:AL445566; PID:gl4089434; F
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPV_0210
A;Genetic code: SGC3
C;Superfamily: dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase homology
C;Keywords: redox-active disulfide
F;41-46/Disulfide bonds: redox-active #status predicted

Query Match 69.4%; Score 34; DB 2; Length 455;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VREQVMKY 10
|||:|
Db 214 VREEVLKY 221

RESULT 17
T17590
probable glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Chlorella
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17590
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17590
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-595 <GRA>

A;Cross-references: UNIPROT:Q84421; UNIPARC:UPI00000F39A6; EMBL:U42580; NID:g4028896; PID:PI
C;Genetics:
A;Note: A100R
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-595/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F;2/Active site: Cys #status predicted

Query Match 69.4%; Score 34; DB 2; Length 595;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VREQVMKY 10
|||:|
Db 378 VRESIMKY 385

RESULT 18
S71091
acetyl-CoA carboxylase (EC 6.4.1.2) beta - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S71091
R;Widmer, J.; Passihi, K.S.; Schlichter, S.C.; Wheeler, K.S.; Crute, B.E.; King, N.; Nut Biochem. J. 316, 915-922, 1996
A;Title: Identification of a second human acetyl-CoA carboxylase gene.
A;Reference number: S71091; MUID:96265061; PMID:8670171
A;Accession: S71091
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-786 <WID>
A;Cross-references: UNIPROT:O00763; UNIPARC:UPI000016B177; EMBL:U34591; NID:g1399289; PID:PI
C;Genetics:
A;Gene: ACC-beta
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: ligase

Query Match 69.4%; Score 34; DB 2; Length 786;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|
Db 286 ESVRYVMRY 295

RESULT 19
S54465
YTA12 protein precursor, mitochondrial - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YM9582.14c; protein YMR089c; RCA1 protein
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S54465; S46609; A55358; S48539
R;Gentles, S.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54451
A;Accession: S54465
A;Molecule type: DNA
A;Residues: 1-825 <GEN>
A;Cross-references: UNIPROT:P40341; UNIPARC:UPI0000133438; EMBL:Z49259; NID:g807956; PID
A;Experimental source: strain AB972
R;Schnall, R.; Mannhaupt, G.; Stucka, R.; Tauer, R.; Ehnle, S.; Schwarzlose, C.; Vetter, Yeast 10, 1141-1155, 1994
A;Title: Identification of a set of yeast genes coding for a novel family of putative ATPases
A;Reference number: S46605; MUID:95274317; PMID:7754704
A;Accession: S46609
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-348,'EL',351-825 <SCH>
A;Cross-references: UNIPARC:UPI0000168FB1; EMBL:X81068; NID:g531753; PIDN:CAA56955.1; PID:PI
R;Tzagoloff, A.; Yue, J.; Jang, J.; Paul, M.F.
J. Biol. Chem. 269, 26144-26151, 1994
A;Title: A new member of a family of ATPases is essential for assembly of mitochondrial

A;Experimental source: clone T23G11
C;Genetics:
A;Gene: CESP:T23G11.5
A;Map position: 1
A;Introns: 41/2; 82/3; 115/3; 199/2

Query Match 67.3%; Score 33; DB 2; Length 243;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 231 ESLSKEVMKY 240

RESULT 24
T03942
rRNA N-glycosidase (EC 3.2.2.22) - maize
N;Alternate names: ribosome-inactivating protein
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03942
R;Bass, H.W.; Obrian, G.R.; Boston, R.S.
Plant Physiol. 107, 661-662, 1995
A;Title: Cloning and sequencing of a second ribosome-inactivating protein gene from maize
A;Reference number: Z15149; MUID:95241633; PMID:7724682
A;Accession: T03942
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-278 <BAS>
A;Cross-references: UNIPROT:Q41851; UNIPARC:UPI000009D113; EMBL:L26305; NID:g496163; PID
A;Experimental source: strain W64A; leaf
C;Genetics:
A;Gene: RIP2
C;Function:
A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: glycosidase; hydrolase
F;27-277/Domain: rRNA N-glycosidase homology <RNG>

Query Match 67.3%; Score 33; DB 2; Length 278;
Best Local Similarity 66.7%; Pred. No. 82;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVREQVMKY 10
|:|:|:|:
Db 43 SVRKEVIKY 51

RESULT 25
A84341
hypothetical protein Vngl900c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C;Accession: A84341
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84341
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-282 <STO>
A;Cross-references: UNIPROT:Q9HNX7; UNIPARC:UPI00001320AD; GB:AE004437; NID:g10581342; F
C;Genetics:
A;Gene: VNG1900C
C;Superfamily: ATP-NAD(H) kinase

Query Match 67.3%; Score 33; DB 2; Length 282;
Best Local Similarity 60.0%; Pred. No. 83;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 103 EAVREEVRY 112

RESULT 26
G82581
GTP binding protein XF2247 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2004
C;Accession: G82581
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <SIM>
A;Cross-references: UNIPROT:Q9PB97; UNIPARC:UPI000012A10D; GB:AE004037; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2247
C;Superfamily: GTP-binding protein era homolog, bacteria type; translation elongation fa

Query Match 67.3%; Score 33; DB 2; Length 298;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
|:|:|:|:
Db 197 ELVREQVMR 205

RESULT 27
H69486
conserved hypothetical protein AF1897 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69486
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69486
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-355 <KLE>
A;Cross-references: UNIPROT:O28382; UNIPARC:UPI0000056B64; GB:AE000972; GB:AE000782; NID

Query Match 67.3%; Score 33; DB 2; Length 355;
Best Local Similarity 70.0%; Pred. No. 1e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||:||

Db 43 ESVREIIIEKY 52

RESULT 28
T43289
fructose-bisphosphate aldolase (EC 4.1.2.13) - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43289
R;Mutoh, N.; Hayashi, Y.
Biochim. Biophys. Acta 1183, 550-552, 1994
A;Title: Molecular cloning and nucleotide sequencing of Schizosaccharomyces pombe homolo
A;Reference number: Z22395; MUID:94114559; PMID:8286404
A;Accession: T43289
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-358 <MUT>
A;Cross-references: UNIPROT:P36580; UNIPARC:UPI0000169149; EMBL:D17415; NID:g398158; PID
C;Superfamily: fructose-bisphosphate aldolase II
C;Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis; pentose ph

Query Match 67.3%; Score 33; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||:||

Db 297 EGVRDYVLKY 306

RESULT 29
T39798
fructose-bisphosphate aldolase (EC 4.1.2.13) - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39798
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21880
A;Accession: T39798
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-358 <MCD>
A;Cross-references: UNIPROT:P36580; UNIPARC:UPI0000125829; EMBL:AL109731; PIDN:CAB52034.
A;Experimental source: strain 972h-; cosmid c19C2
C;Genetics:
A;Gene: SPDB:SPBC19C2.07
A;Map position: 2
C;Superfamily: fructose-bisphosphate aldolase II
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 67.3%; Score 33; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||:||

Db 297 EGVRDYVLKY 306

RESULT 30
AE2092
hypothetical protein all2292 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2092
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-392 <KUR>
A;Cross-references: UNIPROT:Q8YUP3; UNIPARC:UPI00000CE3B3; GB:BA000019; PIDN:BA873991.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2292

Query Match 67.3%; Score 33; DB 2; Length 392;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||:||

Db 156 QSVREEILHY 165

RESULT 31
T12450
hypothetical protein DKFZp564N1916.1 - human
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: T12450
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17523
A;Accession: T12450
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-431 <KOE>
A;Cross-references: UNIPROT:P55010; UNIPARC:UPI000016AC57; EMBL:AL080102
A;Experimental source: fetal brain; clone DKFZp564N1916
C;Genetics:
A;Note: DKFZp564N1916.1

Query Match 67.3%; Score 33; DB 2; Length 431;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||||:||

Db 287 EKTREQIKEY 296

RESULT 32
F86841
iron-binding oxidase subunit [imported] - Lactococcus lactis subsp. lactis (strain IL14;
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86841
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrl
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: F86841
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <STO>
A;Cross-references: UNIPROT:Q9CEV0; UNIPARC:UPI00000C6AEF; GB:AE005176; PID:gl2724753;
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yriC
C;Superfamily: conserved hypothetical iron-sulfur protein HP0138; ferredoxin 2 [4Fe-4S]

Query Match 67.3%; Score 33; DB 2; Length 490;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10

Db 56 ESIRQHTLKY 65 ||:|: :||

RESULT 33
E84667
probable cytochrome P450 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: E84667
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: E84667
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-514 <STO>
A/Cross-references: UNIPROT:Q9ZVD7; UNIPARC:UPI00000A2D48; GB:AE002093; NID:G3885330; PI
C/Genetics:
A/Gene: At2g27000
A/Map position: 2
C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

Query Match 67.3%; Score 33; DB 2; Length 514;
Best Local Similarity 40.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
:::|:|:|

Db 433 DAIKEVLKY 442

RESULT 34
A45249
alpha-glucosidase (EC 3.2.1.20) MAL62 - yeast (Candida albicans)
N/Alternate names: alpha-glucosidase MAL6S; maltase
C/Species: Candida albicans
C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
C/Accession: A45249
R/Geber, A.; Williamson, P.R.; Rex, J.H.; Sweeney, E.C.; Bennett, J.E.
J. Bacteriol. 174, 6992-6996, 1992
A/Title: Cloning and characterization of a Candida albicans maltase gene involved in suc
A/Reference number: A45249; MUID:93015763; PMID:1400249
A/Accession: A45249
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-570 <GEB>
A/Cross-references: UNIPARC:UPI0000175AE1; GB:M94674; NID:G170881; PIDN:AAA34350.1; PID:
A/Note: sequence extracted from NCBI backbone (NCBIP:116994)
C/Superfamily: alpha-glucosidase; alpha-amylase core homology
C/Keywords: glycosidase; hydrolase
F;173-341/Domain: alpha-amylase core homology <AMY>

Query Match 67.3%; Score 33; DB 2; Length 570;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVREQVMKY 10
| | | | :|

Db 267 STREQALKY 275

RESULT 35
T41371
hypothetical protein SPCC4G3.05c - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T41371
R/Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998

A/Reference number: Z21918
A/Accession: T41371
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-572 <WOO>
A/Cross-references: UNIPROT:P87231; UNIPARC:UPI000006B75A; EMBL:Z97052; PIDN:CAB09772.1;
A/Experimental source: strain 972h-; cosmid c4G3
C/Genetics:
A/Gene: SPDB:SPCC4G3.05C
A/Map position: 3

Query Match 67.3%; Score 33; DB 2; Length 572;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
||:| | | | :|

Db 461 ESLREQLLK 469

RESULT 36
A48077
myb-related protein REB1 - yeast (Kluyveromyces marxianus var. lactis)
C/Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C/Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A48077
R/Morrow, B.E.; Ju, Q.; Warner, J.R.
Mol. Cell. Biol. 13, 1173-1182, 1993
A/Title: A bipartite DNA-binding domain in yeast Reb1p.
A/Reference number: A48077; MUID:93140755; PMID:8423784
A/Accession: A48077
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-595 <MOR>
A/Cross-references: UNIPROT:Q05950; UNIPARC:UPI00001334B6; GB:I03789; NID:G173313; PIDN:
A/Note: sequence extracted from NCBI backbone (NCBIN:123574, NCBIP:123576)
C/Superfamily: myb-related protein REB1; myb DNA-binding repeat homology
C/Keywords: DNA binding; nucleus; transcription regulation
F;333-382/Domain: myb DNA-binding repeat homology <MYB>

Query Match 67.3%; Score 33; DB 1; Length 595;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | |

Db 9 ESVEEAVFKY 18

RESULT 37
A97179
ATP-dependent exonuclease synthesis protein AddB (superfamily I helicase) [imported] - C
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: A97179
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: A97179
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1153 <KUR>
A/Cross-references: UNIPROT:Q97GV2; UNIPARC:UPI00000D7548; GB:AE001437; PIDN:AAK80220.1;
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC263
C/Superfamily: Bacillus subtilis ATP-dependent deoxyribonuclease chain B

Query Match 67.3%; Score 33; DB 2; Length 1153;
Best Local Similarity 55.6%; Pred. No. 3.4e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
|:|:|:
Db 1010 EQIREIIMK 1018

RESULT 38
T23429
hypothetical protein K07G5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23429
R;McMurray, A.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19739
A;Accession: T23429
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1202 <WIL>
A;Cross-references: UNIPROT:Q21303; UNIPARC:UPI0000078E43; EMBL:Z71264; PIDN:CAA95830.1;
A;Experimental source: clone K07G5
C;Genetics:
A;Gene: CESP:K07G5.3
A;Map position: 1
A;Introns: 13/1; 104/3; 196/3; 260/2; 282/3; 307/2; 361/2; 410/2; 535/2; 603/2; 688/1; 7
C;Superfamily: Caenorhabditis elegans hypothetical protein K07G5.3

Query Match 67.3%; Score 33; DB 2; Length 1202;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 67 ESIKEHLMKF 76

RESULT 39
A30063
dihydropyridine receptor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A30063
R;Tanabe, T.; Takeshima, H.; Mikami, A.; Flockerzi, V.; Takahashi, H.; Kangawa, K.; Kojima
Nature 328, 313-318, 1987
A;Title: Primary structure of the receptor for calcium channel blockers from skeletal mu
A;Reference number: A30063; MUID:87258269; PMID:3037387
A;Accession: A30063
A;Molecule type: mRNA
A;Residues: 1-1873 <TAN>
A;Cross-references: UNIPROT:P07293; UNIPARC:UPI000005031B; GB:X05921; NID:gl547; PIDN:CA
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C;Keywords: phosphoprotein; transmembrane protein

Query Match 67.3%; Score 33; DB 2; Length 1873;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 825 ESVRNQILGY 834

RESULT 40
S69116
fibrinogen gamma chain - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Mar-1998 #sequence_revision 24-Apr-1998 #text_change 02-Jul-1998
C;Accession: S69116
R;Litvinovich, S.V.; Henschen, A.H.; Krieglststein, K.G.; Ingham, K.C.; Medved, L.V.
Eur. J. Biochem. 229, 605-614, 1995
A;Title: Structural and functional characterization of proteolytic fragments derived fro
A;Reference number: S69114; MUID:95278210; PMID:7758453
A;Accession: S69116

A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-92 <LIT>
A;Cross-references: UNIPARC:UPI00001777E6
C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

Query Match 65.3%; Score 32; DB 2; Length 92;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 26 KSMEEIIMKY 35

RESULT 41
T27724
hypothetical protein ZK1251.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27724
R;McMurray, A.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z20411
A;Accession: T27724
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-138 <WIL>
A;Cross-references: UNIPROT:Q23434; UNIPARC:UPI00000793E7; EMBL:Z68222; PIDN:CAA92503.1;
A;Experimental source: clone ZK1251
C;Genetics:
A;Gene: CESP:ZK1251.7
A;Map position: 4
C;Superfamily: Caenorhabditis elegans hypothetical protein ZK1251.7

Query Match 65.3%; Score 32; DB 2; Length 138;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 25 ENIRQQATKY 34

RESULT 42
E97310
uncharacterized conserved domain seen in the bacterial SpOT CAC3340 [imported] - Clostr
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97310
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97310
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <KUR>
A;Cross-references: UNIPROT:Q97DY0; UNIPARC:UPI00000D46F2; GB:AE001437; PIDN:AAK81272.1
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3340
C;Superfamily: GTP pyrophosphokinase related protein

Query Match 65.3%; Score 32; DB 2; Length 217;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 20 EEMQEMIMKY 29

RESULT 43
H64321
hypothetical protein MJ0175 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: H64321
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: H64321
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-222 <BUL>
A;Cross-references: UNIPROT:Q57639; UNIPARC:UPI00001393DD; GB:U67474; GB:L77117; NID:g15
C;Genetics:
A;Map position: REV178140-177472
A;Start codon: GTG

Query Match 65.3%; Score 32; DB 2; Length 222;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
:|||||:
Db 32 DSVREILK 40

RESULT 44
S45880
SCO1 protein homolog SCO2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR024w; hypothetical protein YBR0308
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: S45880; S46555
R;Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45875
A;Accession: S45880
A;Molecule type: DNA
A;Residues: 1-301 <GRI>
A;Cross-references: UNIPROT:P38072; UNIPARC:UPI000013566E; EMBL:Z35893; NID:g536231; PID
A;Experimental source: strain S289C
R;Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
Yeast 10(Suppl.A), S75-S80, 1994
A;Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A;Reference number: S46551; MUID:94378725; PMID:8091864
A;Accession: S46555
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <SMI>
A;Cross-references: UNIPARC:UPI000013566E; EMBL:X76078; NID:g498748; PID:g498753
A;Experimental source: strain S289C
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Gene: SGD:SCO2
A;Cross-references: SGD:S0000228; MIPS:YBR024w
A;Map position: 2R
C;Keywords: mitochondrion; transmembrane protein

Query Match 65.3%; Score 32; DB 2; Length 301;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
:|||||:
Db 273 EKIREIQAY 282

RESULT 45
S00909
replication initiation protein - Staphylococcus aureus plasmid pC223
C;Species: Staphylococcus aureus
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S00909
R;Ehret, M.; Matzura, H.
Nucleic Acids Res. 16, 2045-2062, 1988
A;Title: Replication control of the Staphylococcus aureus chloramphenicol resistance pla
A;Reference number: S00909; MUID:88189801; PMID:3128771
A;Accession: S00909
A;Molecule type: DNA
A;Residues: 1-314 <EHR>
A;Cross-references: UNIPROT:P14490; UNIPARC:UPI00001336A6; EMBL:X07371; NID:g46635; PIDN
A;Note: the authors translated the codon GAT for residue 64 as Lys, GCA for residue 174
C;Genetics:
A;Gene: repM
A;Genome: plasmid
C;Superfamily: repC protein

Query Match 65.3%; Score 32; DB 2; Length 314;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|:
Db 243 ESLKEQAMVY 252

RESULT 46
A30480
repJ protein - Staphylococcus aureus plasmid pC223
N;Alternate names: replication initiation protein
C;Species: Staphylococcus aureus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A30480
R;Projan, S.J.; Novick, R.
Plasmid 19, 203-221, 1988
A;Title: Comparative analysis of five related staphylococcal plasmids.
A;Reference number: JT0372; MUID:89161003; PMID:2852816
A;Accession: A30480
A;Molecule type: DNA
A;Residues: 1-314 <PRO>
A;Cross-references: UNIPROT:Q52239; UNIPARC:UPI00001783DB
A;Note: the authors translated the codon tta for residue 123 as Ile
C;Genetics:
A;Gene: repJ
A;Genome: plasmid
C;Superfamily: repC protein
F;238-282/Domain: DNA binding #status predicted <DNB>

Query Match 65.3%; Score 32; DB 2; Length 314;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|:
Db 243 ESLKEQAMVY 252

RESULT 47
T49224
SRG1-like protein - Arabidopsis thaliana
N;Alternate names: protein F27H5.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49224
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25018
A;Accession: T49224
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-316 <RIE>
A;Cross-references: UNIPROT:Q9LY36; UNIPARC:UPI00000A82E0; EMBL:AL163852; GSPDB:GN00061;
A;Experimental source: cultivar Columbia; BAC clone F27H5
C;Genetics:
A;Gene: ATSP:F27H5.80
A;Map position: 3
A;Introns: 84/3; 126/2
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 65.3%; Score 32; DB 2; Length 316;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
: |||: ||
Db 123 EPVREKVGKY 132

RESULT 48
T04309
acidic ribosomal protein P0 - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04309
R;Hiwara, Y.; Umeda, M.; Hara, C.; Toriyama, K.; Uchimiya, H.
Plant Physiol. 105, 753-754, 1994
A;Title: Nucleotide sequence of a rice acidic ribosomal phosphoprotein P0 cDNA.
A;Reference number: Z15273; MUID:94345013; PMID:8066137
A;Accession: T04309
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-319 <HIH>
A;Cross-references: UNIPROT:P41095; UNIPARC:UPI000013422D; EMBL:D21130; NID:G415316; PID
A;Experimental source: anther
C;Genetics:
A;Gene: YK704
C;Superfamily: rat acidic ribosomal protein P0

Query Match 65.3%; Score 32; DB 2; Length 319;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
: |||: ||
Db 99 KEVREEVAKY 108

RESULT 49
T03944
acidic ribosomal protein P0 - maize
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03944
R;Cordts, S.; Loerz, H.; Dresselhaus, T.
submitted to the EMBL Data Library, September 1996
A;Description: A transcript encoding the large subunit acidic ribosomal protein P0 is s
A;Reference number: Z15151
A;Accession: T03944
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-319 <COR>
A;Cross-references: UNIPROT:Q24573; UNIPARC:UPI000016E078; EMBL:Y07959; PIDN:CAA69256.1
A;Experimental source: note inbred line A188
C;Genetics:
A;Note: RP-P0
C;Superfamily: rat acidic ribosomal protein P0
C;Keywords: protein biosynthesis; ribosome; RNA binding

Query Match 65.3%; Score 32; DB 2; Length 319;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
: |||: ||
Db 99 KEVREEVAKY 108

RESULT 50
T22701
hypothetical protein F55B12.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22701
R;Sims, M.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19602
A;Accession: T22701
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-350 <WIL>
A;Cross-references: UNIPROT:Q93797; UNIPARC:UPI000007E666; EMBL:Z79757; PIDN:CAB02127.1;
A;Experimental source: clone F55B12
C;Genetics:
A;Gene: CESP:F55B12.6
A;Map position: 5
A;Introns: 64/3; 94/1; 180/2; 227/3; 271/3; 331/1

Query Match 65.3%; Score 32; DB 2; Length 350;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
: |||: ||
Db 177 ELIRESVMKH 186

Search completed: May 12, 2006, 10:52:04
Job time : 25.2368 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-38
Perfect score: 49
Sequence: 1 ESVREQVMKY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	83.7	458	Q4S6F3_TETNG	Q486f3 tetraodon n
2	40	81.6	1164	Q9ZJN3_HELPJ	Q9zjn3 helicobacte
3	39	79.6	1064	Q4SFR7_TETNG	Q4sfr7 tetraodon n
4	38	77.6	291	ERA_MYCPN	P75210 mycoplasma
5	38	77.6	592	Q52A99_MAGGR	Q52a99 magnaporthe
6	38	77.6	593	Q5CP31_CRYHO	Q5cp31 cryptospori
7	38	77.6	593	Q5CYP8_CRYPV	Q5cyp8 cryptospori
8	38	77.6	2448	Q6JIZ0_MOUSE	Q6jiz0 mus musculu
9	37	75.5	394	Q9LJZ1_ARATH	Q9ljz1 arabidopsis
10	37	75.5	429	IF5_MOUSE	P59325 mus musculu
11	37	75.5	429	IF5_RAT	Q07205 rattus norv
12	37	75.5	430	Q5ZIE0_CHICK	Q5zie0 gallus gall
13	37	75.5	431	IF5_HUMAN	P55010 homo sapien
14	37	75.5	431	Q53XB3_HUMAN	Q53xb3 homo sapien
15	37	75.5	431	Q5R4L0_PONPY	Q5r4l0 pongo pygma
16	37	75.5	892	Q7QPI9_GIALA	Q7qpi9 giardia lam
17	37	75.5	1132	Q8I5C3_PLAF7	Q8i5c3 plasmodium
18	37	73.5	150	RS13_YEAST	P05756 saccharomyc
19	36	73.5	151	Q6FT65_CANGA	Q6ft65 candida gla
20	36	73.5	160	YBIA_ECOLI	P30176 escherichia
21	36	73.5	160	Q83L08_SHIFL	Q83lu8 shigella fl
22	36	73.5	187	Q50HZ9_BOVIN	Q50hz9 bos taurus
23	36	73.5	193	Q4P3L0_USTMA	Q4p3l0 ustilago ma
24	36	73.5	242	Q62DC6_BURMA	Q62dc6 burkholderi
25	36	73.5	245	Q63QH4_BURPS	Q63qh4 burkholderi
26	36	73.5	418	Q6A7W0_PROAC	Q6a7w0 propionibac
27	36	73.5	499	Q9FH67_ARATH	Q9fh67 arabidopsis
28	36	73.5	523	Q9LI99_ARATH	Q9li99 arabidopsis
29	36	73.5	671	Q529X0_MAGGR	Q529x0 magnaporthe
30	36	73.5	829	Q51V60_MAGGR	Q51v60 magnaporthe
31	36	73.5	970	Q5CWC1_CRYPV	Q5cwc1 cryptospori

32	36	73.5	998	Q59FY4_HUMAN	Q59fy4 homo sapien
33	36	73.5	1182	Q4SCU4_TETNG	Q4scu4 tetraodon n
34	36	73.5	1233	Q5F363_CHICK	Q5f363 gallus gall
35	36	73.5	1234	JARD2_MOUSE	Q62315 mus musculu
36	36	73.5	1499	Q86LZ8_SCHMA	Q86lzl8 schistosoma
37	36	73.5	1847	Q8IDM0_PLAF7	Q8idm0 plasmodium
38	36	73.5	2324	COAC_CHICK	P11029 gallus gall
39	36	73.5	2345	COA1_RAT	P11497 rattus norv
40	36	73.5	2345	Q5SWU9_MOUSE	Q5swu9 mus musculu
41	36	73.5	2345	Q6JIZ1_MOUSE	Q6jiz1 mus musculu
42	36	73.5	2346	COA1_BOVIN	Q9ttts3 bos taurus
43	36	73.5	2346	COA1_HUMAN	Q13085 homo sapien
44	36	73.5	2346	COA1_SHEEP	Q28559 ovis aries
45	36	73.5	2346	Q7Z561_HUMAN	Q7z561 homo sapien
46	36	73.5	2346	Q6XDA8_HUMAN	Q6xda8 homo sapien
47	35	71.4	74	Q9BS18_HUMAN	Q9bs18 homo sapien
48	35	71.4	74	Q5RBV4_PONPY	Q5rbv4 pongo pygma
49	35	71.4	74	Q8R034_MOUSE	Q8r034 mus musculu
50	35	71.4	74	Q4SKC6_TETNG	Q4skc6 tetraodon n
51	35	71.4	82	Q5XZ64_BORGA	Q5xz64 borrelia ga
52	35	71.4	127	Q87310_BORBU	Q87310 borrelia bu
53	35	71.4	140	Q4RCQ3_TETNG	Q4rcq3 tetraodon n
54	35	71.4	193	Q6MT64_MYCMS	Q6mt64 mycoplasma
55	35	71.4	193	Q5ZKG7_CHICK	Q5zkg7 gallus gall
56	35	71.4	197	Q8MVN6_9ASCI	Q8mvn6 boltenia vi
57	35	71.4	215	Q9HKP2_THEAC	Q9hkp2 thermoplasm
58	35	71.4	217	Q97B50_THEVO	Q97b50 thermoplasm
59	35	71.4	248	Q5ZM54_CHICK	Q5zm54 gallus gall
60	35	71.4	424	Q9R2Y2_BORBU	Q9r2y2 borrelia bu
61	35	71.4	424	Q9S0C6_BORBU	Q9s0c6 borrelia bu
62	35	71.4	424	Q9S016_BORBU	Q9s0i6 borrelia bu
63	35	71.4	424	Q9S049_BORBU	Q9s049 borrelia bu
64	35	71.4	429	Q9L9P1_BORBU	Q9l9p1 borrelia bu
65	35	71.4	429	Q7ZWA9_BRARE	Q7zwa9 brachydanio
66	35	71.4	431	Q9S000_BORBU	Q9s000 borrelia bu
67	35	71.4	431	Q9S0F6_BORBU	Q9s0f6 borrelia bu
68	35	71.4	436	Q5XYW7_BORGA	Q5xyw7 borrelia ga
69	35	71.4	444	Q9S093_BORBU	Q9s093 borrelia bu
70	35	71.4	446	Q5ZMN6_CHICK	Q5zmn6 gallus gall
71	35	71.4	486	Q9WZ51_THEMEA	Q9wz51 thermotoga
72	35	71.4	526	Q9LIQ8_ARATH	Q9lig8 arabidopsis
73	35	71.4	596	Q76DQ7_9PHYC	Q76dq7 chlorella v
74	35	71.4	653	Q6FRB3_CANGA	Q6frb3 candida gla
75	35	71.4	698	Q51X41_MAGGR	Q51x41 magnaporthe
76	35	71.4	753	Q4N7W0_THEPA	Q4n7w0 theileria p
77	35	71.4	810	REB1_YEAST	P21538 saccharomyc
78	35	71.4	869	Q60SF7_CAEBR	Q60sf7 caenorhabdi
79	35	71.4	1003	Q8IAC0_HALRO	Q8iac0 halocynthia
80	35	71.4	1271	P94460_BACSU	P94460 bacillus su
81	35	71.4	1272	Q799M0_BACSU	Q799m0 bacillus su
82	35	71.4	1279	Q31827_BACSU	Q31827 bacillus su
83	35	71.4	1317	Q54X69_DICDI	Q54x69 dictyosteli
84	35	71.4	2456	Q70151_RAT	Q70151 rattus norv
85	34	69.4	99	Q6LNQ3_PHOPR	Q6lnq3 photobacter
86	34	69.4	132	Q4XH90_PLACH	Q4xh90 plasmodium
87	34	69.4	159	Q6YD94_9FABA	Q6ydg4 arachis ste
88	34	69.4	159	Q6YDF0_9FABA	Q6ydf0 arachis car
89	34	69.4	175	Q81EH3_BACCR	Q81eh3 bacillus ce
90	34	69.4	175	Q63CF3_BACCCZ	Q63cf3 bacillus ce
91	34	69.4	204	Q8SSK8_ENCCU	Q8ssk8 encephalito
92	34	69.4	212	Q9KRL5_VIBCH	Q9krl5 vibrio chol
93	34	69.4	212	Q8NLY1_CORGL	Q8nly1 corynebacte
94	34	69.4	219	BUN1_DROME	Q24522 drosophila
95	34	69.4	272	Q98QV3_MYCPU	Q98qv3 mycoplasma
96	34	69.4	287	Q48615_9LACT	Q48615 lactococcus
97	34	69.4	357	Q8Z110_SALTI	Q8z110 salmonella
98	34	69.4	383	Q59RJ2_CANAL	Q59rj2 candida alb
99	34	69.4	407	Q5R5T5_PONPY	Q5r5t5 pongo pygma
100	34	69.4	412	Q8QGW4_BRARE	Q8qgw4 brachydanio
101	34	69.4	416	Q6MM69_BDEBA	Q6mm69 bdellovibri
102	34	69.4	450	Q4R7L2_MACFA	Q4r7l2 macaca fasc
103	34	69.4	455	Q98RI8_MYCPU	Q98ri8 mycoplasma
104	34	69.4	481	Q584W0_9TRYP	Q584w0 trypanosoma

105	34	69.4	504	2	Q896Y2_CLOTE	Q896Y2_clostridium	178	33	67.3	461	2	Q5A441_CANAL	Q5a441_candida alb
106	34	69.4	513	2	Q8LEJ8_ARATH	Q8lej8_arabidopsis	179	33	67.3	469	2	Q4ZXX9_PSESY	Q4zxx9_pseudomonas
107	34	69.4	513	2	Q9LJY8_ARATH	Q9ljy8_arabidopsis	180	33	67.3	490	2	Q9CEV0_LACLA	Q9cev0_lactococcus
108	34	69.4	516	2	Q6BU60_DEBHA	Q6bu60_debaryomyce	181	33	67.3	497	2	Q7RLR0_PLAYO	Q7rlr0_plasmodium
109	34	69.4	519	2	Q6A8M6_PROAC	Q6a8m6_propionibac	182	33	67.3	505	2	Q891M1_CLOTE	Q891m1_clostridium
110	34	69.4	573	1	SYE_PYRKO	Q6a8m6_propionibac	183	33	67.3	512	2	Q9XV10_CAEEL	Q9xv10_caenorhabdi
111	34	69.4	595	2	Q84421_CHVP1	Q5jhl6_pyrococcus	184	33	67.3	514	2	Q9ZVD7_ARATH	Q9zvd7_arabidopsis
112	34	69.4	752	2	Q7Q3A1_ANOGA	Q84421_paramecium	185	33	67.3	542	2	Q6WGM5_9PARA	Q6wgm5_mosman vir
113	34	69.4	811	1	PLSB_HAEDU	Q7q3al_anopheles g	186	33	67.3	561	2	Q55VJ1_CRYNE	Q55vj1_cryptococcu
114	34	69.4	825	1	RCA1_YEAST	Q7vni5_aemophilus	187	33	67.3	561	2	Q5KKP8_CRYNE	Q5kqp8_cryptococcu
115	34	69.4	825	2	Q6B1I1_YEAST	P40341_saccharomyc	188	33	67.3	567	2	Q6MQB2_BDEBA	Q6mqb2_bdellovibri
116	34	69.4	834	2	Q8SXR0_DROME	Q6b1l1_saccharomyc	189	33	67.3	569	1	MALT_CANAL	Q02751_candida alb
117	34	69.4	858	2	Q4GL170_HUMAN	Q8sxxr0_drosophila	190	33	67.3	570	2	Q5ACH4_CANAL	Q5ach4_candida alb
118	34	69.4	861	2	Q4HB46_9DEIO	Q4gl170_homo sapien	191	33	67.3	572	1	MUS81_SCHPO	P87231_schizosacch
119	34	69.4	908	2	Q25646_PLABE	Q4hb46_deinococcus	192	33	67.3	595	1	REB1_KLULA	Q05950_kluyveromyc
120	34	69.4	916	1	CADH4_HUMAN	Q25646_plasmodium	193	33	67.3	609	2	Q8JRU0_9BACU	Q8jru0_phthorimaea
121	34	69.4	924	2	Q9NA99_CAEEL	P55283_homo sapien	194	33	67.3	673	2	Q5WXV3_LEGPL	Q5wxv3_legionella
122	34	69.4	940	1	UVRA_VIBVU	Q8dcj3_vibrio vuln	195	33	67.3	673	2	Q5ZWX6_LEGPH	Q5zxw6_legionella
123	34	69.4	940	1	UVRA_VIBVY	Q7mhb5_vibrio vuln	196	33	67.3	683	2	Q9FMN1_ARATH	Q9fmn1_arabidopsis
124	34	69.4	949	2	Q8E8Z1_SHEON	Q8e8z1_shewanella	197	33	67.3	751	2	Q94LH1_ORYSA	Q9fnn1_arabidopsis
125	34	69.4	960	2	Q8JMH5_9VIRU	Q8jmh5_cherry rasp	198	33	67.3	776	2	Q5GAP5_MAIZE	Q5gap5_zea mays (m
126	34	69.4	960	2	Q6EWG8_9VIRU	Q6ewg8_cherry rasp	199	33	67.3	784	2	Q75WV8_BRABE	Q75wv8_branchiosto
127	34	69.4	968	2	Q5P4I8_AZOSE	Q5p4i8_azoarcus sp	200	33	67.3	798	2	Q8LHQ6_ORYSA	Q8lhq6_oryza sativ
128	34	69.4	984	2	Q4Z4G1_PLABE	Q4z4g1_plasmodium	201	33	67.3	817	2	Q7YYH4_CRYPV	Q7yyh4_cryptospori
129	34	69.4	986	2	Q4XZ91_PLACH	Q4xz91_plasmodium	202	33	67.3	836	2	RABE1_HUMAN	Q15276_homo sapien
130	34	69.4	1048	2	Q7RS02_PLAYO	Q7rs02_plasmodium	203	33	67.3	862	1	Q5CX78_CRYPV	Q5cx78_cryptospori
131	34	69.4	1070	2	Q8IB03_PLAF7	Q8ib03_plasmodium	204	33	67.3	896	2	Q7NBT1_MYCGA	Q7nbt1_mycoplasma
132	34	69.4	1689	2	Q59GJ9_HUMAN	Q59gj9_homo sapien	205	33	67.3	921	2	Q8ZJY3_SALT	Q8zjy3_salmonella
133	34	69.4	1838	2	Q8IDZ6_PLAF7	Q8idz6_plasmodium	206	33	67.3	921	2	Q6CL70_KLULA	Q6cl70_kluyveromyc
134	34	69.4	2458	2	Q6KE87_HUMAN	Q6ke87_homo sapien	207	33	67.3	957	2	Q580R3_TRYPA	Q580r3_trypanosoma
135	34	69.4	2458	2	Q6TY48_HUMAN	Q6ty48_homo sapien	208	33	67.3	1012	2	Q6FLN2_CANGA	Q6fln2_candida gla
136	34	69.4	2483	1	COA2_HUMAN	Q00763_homo sapien	209	33	67.3	1030	2	Q97GV2_CLOAB	Q97gv2_clostridium
137	33	67.3	72	2	Q5QQP4_PLARU	Q5qpp4_planktothri	210	33	67.3	1153	2	Q21303_CAEEL	Q21303_caenorhabdi
138	33	67.3	72	2	Q5QQQ6_PLARU	Q5qqq6_planktothri	211	33	67.3	1202	2	Q60T12_CAEER	Q60t12_caenorhabdi
139	33	67.3	72	2	Q5QQQ8_PLARU	Q5qqq8_planktothri	212	33	67.3	1207	2	Q7PDW0_PLAYO	Q7pdw0_plasmodium
140	33	67.3	72	2	Q5QQR2_OSCAG	Q5qqr2_oscillatori	213	33	67.3	1287	2	Q5KHJ6_CRYNE	Q5khj6_cryptococcu
141	33	67.3	72	2	Q5QQR6_OSCAG	Q5qqr6_oscillatori	214	33	67.3	1451	2	Q55UU2_CRYNE	Q55uu2_cryptococcu
142	33	67.3	72	2	Q5QQS0_OSCAG	Q5qqso_oscillatori	215	33	67.3	1462	2	CAC1S_RABIT	P07293_oryctolagus
143	33	67.3	72	2	Q5QQS2_OSCAG	Q5qqs2_oscillatori	216	33	67.3	1873	1	Q4RSU6_TETNG	Q4rsu6_tetraodon n
144	33	67.3	111	2	Q9WMG1_9VIRU	Q9wmgi_topografov	217	33	67.3	2365	2	Q8G986_OSCAG	Q8g986_oscillatori
145	33	67.3	133	2	Q4RJ93_TETNG	Q8rj93_tetraodon n	218	33	67.3	3488	2	Q72UT4_LEPIC	Q72ut4_leptospira
146	33	67.3	136	2	Q872H2_NEUCR	Q872h2_neurospora	219	32	65.3	69	2	Q6LNF8_PHOPR	Q6lnf8_photobacter
147	33	67.3	159	2	Q5JE92_PYRKO	Q5je92_pyrococcus	220	32	65.3	103	2	Q517E3_ENTHI	P38505_entamoeba h
148	33	67.3	163	2	Q704P4_BOVIN	Q704p4_bos taurus	221	32	65.3	134	1	Q23434_CAEEL	Q23434_caenorhabdi
149	33	67.3	179	2	Q97W67_SULSO	Q97w67_sulfolobus	222	32	65.3	138	2	Q7N9G4_PHOLL	Q7n9g4_photorhabdu
150	33	67.3	219	2	Q4H9X6_9DEIO	Q4h9x6_deinococcus	223	32	65.3	165	2	Q9P4B4_LETVU	Q9p4b4_letharia vu
151	33	67.3	228	2	Q65GC2_BACLD	Q65gc2_bacillus li	224	32	65.3	170	2	Q8XL92_CLOPE	Q8xl92_clostridium
152	33	67.3	247	2	Q7RYB7_NEUCR	Q7ryb7_neurospora	225	32	65.3	170	2	Q848T4_ENTFC	Q848t4_osterococcu
153	33	67.3	270	2	Q8A0A3_BACTN	Q8a0a3_bacteroides	226	32	65.3	182	2	Q5M8K8_XENTR	Q5m8k8_xenopus tro
154	33	67.3	273	2	Q4Z306_PLABE	Q4z306_plasmodium	227	32	65.3	184	2	Q5C190_SCHJA	Q5c190_schistosoma
155	33	67.3	278	2	Q41851_MAIZE	Q41851_zea mays (m	228	32	65.3	193	2	Q678E9_9VIRU	Q678e9_lymphocysti
156	33	67.3	280	2	Q5LGU3_BACFN	Q5lgu3_bacteroides	229	32	65.3	197	2	Q6Q912_9GAMM	Q6q912_uncultured
157	33	67.3	280	2	Q64XP6_BACFR	Q64xp6_bacteroides	230	32	65.3	198	2	Q6P2S7_HUMAN	Q6p2s7_homo sapien
158	33	67.3	282	1	PPNK_HALSA	Q9hnx7_halobacteri	231	32	65.3	200	2	Q8LNX8_ZINEL	Q8lnx8_zinnia eleg
159	33	67.3	298	1	ERA_XYLFA	Q9pb97_xylella fas	232	32	65.3	203	2	Q5T829_HUMAN	Q5t829_homo sapien
160	33	67.3	298	1	ERA_XYLFT	Q87c05_xylella fas	233	32	65.3	205	2	Q5E243_VIBF1	Q5e243_vibrio fisc
161	33	67.3	304	2	Q6CFF7_KLULA	Q6cfp7_kluyveromyc	234	32	65.3	214	2	Q97DY0_CLOAB	Q97dy0_clostridium
162	33	67.3	308	2	Q74LN8_LACJO	Q74ln8_lactobacill	235	32	65.3	222	1	Y175_METJA	Q57639_methanococc
163	33	67.3	354	2	Q6XM89_XENLA	Q6xm89_xenopus lae	236	32	65.3	224	2	Q993S6_9VIRU	Q993s6_banana mild
164	33	67.3	354	2	Q7T101_XENLA	Q7t101_xenopus lae	237	32	65.3	227	2	Q516A0_ENTHI	Q516a0_entamoeba h
165	33	67.3	354	2	Q4QR41_XENLA	Q4qr41_xenopus lae	238	32	65.3	227	2	Q8DKV2_ZYNEL	Q8dkv2_synecococc
166	33	67.3	355	2	Q28382_ARCFU	Q28382_archaeoglob	239	32	65.3	236	2	Q600A4_MYCHY	Q600a4_mycoplasma
167	33	67.3	358	1	ALF_SCHPO	P36580_schizosacch	240	32	65.3	240	2	Q5YU16_NOCFA	Q5yul6_nocardia fa
168	33	67.3	366	2	Q50PH3_ENTHI	Q50ph3_entamoeba h	241	32	65.3	247	2	Q86D97_SCHMA	Q86d97_schistosoma
169	33	67.3	367	2	Q50VE8_ENTHI	Q50ve8_entamoeba h	242	32	65.3	251	2	Q84VC9_ORYSA	Q84vc9_oryza sativ
170	33	67.3	377	2	Q6DG41_BRARE	Q6dg41_brachydanio	243	32	65.3	252	2	Q4X1T0_ASPFU	Q4x1t0_aspergillus
171	33	67.3	384	2	Q54I38_DICDI	Q54i38_dictyosteli	244	32	65.3	260	2	Q7ZYS5_XENLA	Q7zys5_xenopus lae
172	33	67.3	392	2	Q8YUP3_ANASP	Q8yup3_anabaena sp	245	32	65.3	260	2	Q7SZA1_XENLA	Q7sza1_xenopus lae
173	33	67.3	398	2	Q5QVZ0_IDILO	Q5qvz0_idiomarina	246	32	65.3	261	2	Q6DDC5_XENTR	Q6ddc5_xenopus tro
174	33	67.3	400	1	ACKA_PHOLL	Q7n2i1_photorhabdu	247	32	65.3	264	2	Q8AW09_BRARE	Q8aw09_brachydanio
175	33	67.3	413	2	Q8TMS1_METAC	Q8tms1_methanosarc	248	32	65.3	266	2	Q51SN5_MACFA	Q51sn5_macaca fasc
176	33	67.3	431	2	Q61BU0_HUMAN	Q6ibu0_homo sapien	249	32	65.3				
177	33	67.3	444	2	Q8A735_BACTN	Q8a735_bacteroides	250	32	65.3				

251	32	65.3	275	2	Q4L7A8_STAHLJ	Q417a8	staphylococ
252	32	65.3	286	2	Q7ZZ43_BRARE	Q7zz43	brachydanio
253	32	65.3	301	1	SCO2_YEAST	P38072	saccharomyc
254	32	65.3	303	2	Q74IY9_LACJO	Q74iy9	lactobacilli
255	32	65.3	304	2	Q4HN33_CAMLA	Q4hn33	campylobact
256	32	65.3	314	1	REPM_STAAL	P14490	staphylococ
257	32	65.3	314	2	Q52239_9ZZZZ	Q52239	plasmid pc2
258	32	65.3	314	2	Q6V0Y3_STAAL	Q6v0y3	staphylococ
259	32	65.3	316	2	Q9LY36_ARATH	Q9ly36	arabidopsi
260	32	65.3	318	1	RLA0_MAIZE	O24573	zea mays (m
261	32	65.3	319	1	RLA0_ORYSA	P41095	oryza sativ
262	32	65.3	319	2	Q7EZR2_ORYSA	Q7ezr2	oryza sativ
263	32	65.3	330	2	Q5KVB0_GEOKA	Q5kvb0	geobacillus
264	32	65.3	337	2	Q54EC9_DICDI	Q54ec9	dictyosteli
265	32	65.3	339	2	Q6CH22_YARLI	Q6ch22	yarrowia li
266	32	65.3	349	2	Q5GHF0_LACRH	Q5ghf0	lactobacilli
267	32	65.3	350	2	Q93797_CAEEL	Q93797	caenorhabdi
268	32	65.3	350	2	Q8PER8_XANAC	Q8per8	xanthomonas
269	32	65.3	354	2	Q9QZA9_MOUSE	Q9qza9	mus musculu
270	32	65.3	359	2	Q5I0L6_RAT	Q5i0l6	rattus norv
271	32	65.3	359	2	Q91VY6_MOUSE	Q91vy6	mus musculu
272	32	65.3	375	2	Q8F3S2_LEPIN	Q8f3s2	leptospi
273	32	65.3	375	2	Q72RX4_LEPIC	Q72rx4	leptospi
274	32	65.3	376	1	OGG1_YEAST	P53397	saccharomyc
275	32	65.3	392	2	Q6GN39_XENLA	Q6gn39	xenopus lae
276	32	65.3	394	2	Q64CZ1_9ARCH	Q64cz1	uncultured
277	32	65.3	403	1	METK_XANAC	Q8pp75	xanthomonas
278	32	65.3	403	1	METK_XANCP	Q8pch3	xanthomonas
279	32	65.3	403	1	METK_XYLFA	Q9pgb0	xylella fas
280	32	65.3	403	1	METK_XYLFT	Q87ay6	xylella fas
281	32	65.3	403	2	Q5GW76_XANOR	Q5gw76	xanthomonas
282	32	65.3	403	2	Q4UR08_XANCP	Q4ur08	xanthomonas
283	32	65.3	403	2	Q88GQ1_PSEPK	Q88gq1	pseudomonas
284	32	65.3	403	2	Q97LQ5_CLOAB	Q97lq5	clostridium
285	32	65.3	410	2	Q9YCU8_AERPE	Q9ycu8	aeropyrum p
286	32	65.3	410	2	P96482_STRPN	P96482	streptococ
287	32	65.3	410	2	Q4K386_STRPN	Q4k386	streptococ
288	32	65.3	412	2	Q9X9A9_STRPN	Q9x9a9	streptococ
289	32	65.3	412	2	Q9AH99_STRPN	Q9ah99	streptococ
290	32	65.3	428	2	Q4JAY2_SULAC	Q4jay2	sulfolobus
291	32	65.3	433	2	Q4NCL8_9MICC	Q4nc18	arthrobacte
292	32	65.3	444	1	FIBG_BOVIN	P12799	bos taurus
293	32	65.3	445	2	Q97J59_CLOAB	Q97j59	clostridium
294	32	65.3	445	2	Q9KDG8_BACHD	Q9kdg8	bacillus ha
295	32	65.3	451	2	Q83DH6_COXBU	Q83dh6	coxiella bu
296	32	65.3	460	2	Q6N8G4_RHOPA	Q6n8g4	rhodopseudo
297	32	65.3	468	2	Q9LCC6_9BACI	Q9lcc6	bacillus sp
298	32	65.3	470	2	Q4H579_9DEIO	Q4h579	deinococcus
299	32	65.3	476	2	Q6BUY1_DEBHA	Q6buy1	debaryomyce
300	32	65.3	488	2	Q4RWX7_TETNG	Q4rxw7	tetraodon n
301	32	65.3	489	1	Y092_METJA	Q57557	methanococ
302	32	65.3	496	2	Q4QCG7_LEIMA	Q4qcg7	leishmania
303	32	65.3	501	2	Q9WY3_THEMEA	Q9wyy3	thermotoga
304	32	65.3	526	2	Q68F97_XENTR	Q68f97	xenopus tro
305	32	65.3	530	1	WC2_NEUCR	P78714	neurospora
306	32	65.3	530	2	Q7M8P0_WOLSU	Q7m8p0	wolinella s
307	32	65.3	531	2	Q67J74_SYMTH	Q67j74	symbiobacte
308	32	65.3	538	2	Q7RVJ5_NEUCR	Q7rvj5	neurospora
309	32	65.3	548	2	Q7PYS5_ANOGA	Q7py55	anopheles g
310	32	65.3	550	2	Q60WT6_CABBR	Q60wt6	caenorhabdi
311	32	65.3	550	2	O22880_ARATH	O22880	arabidopsis
312	32	65.3	570	2	Q12457_YEAST	Q12457	saccharomyc
313	32	65.3	573	2	Q4SUC2_TETNG	Q4suc2	tetraodon n
314	32	65.3	578	2	Q60J20_CAEER	Q60j20	caenorhabdi
315	32	65.3	589	2	Q501D5_ARATH	Q50ld5	arabidopsis
316	32	65.3	595	2	Q53MB3_ORYSA	Q53mb3	oryza sativ
317	32	65.3	658	2	Q9AT32_DAUCA	Q9at32	daucus caro
318	32	65.3	678	1	VPS1_SCHPO	Q9ur25	schizosacch
319	32	65.3	685	2	Q51JQ7_MAGGR	Q51jq7	magnaporthe
320	32	65.3	711	2	Q8I4I1_CAEEL	Q8i4i1	caenorhabdi
321	32	65.3	714	2	Q7SX85_BRARE	Q7sx85	brachydanio
322	32	65.3	716	2	Q8C449_MOUSE	Q8c449	mus musculu
323	32	65.3	721	2	Q87A00_XYLFT	Q87a00	xylella fas

324	32	65.3	721	2	Q9PA43_XYLFA	Q9pa43	xylella fas
325	32	65.3	732	2	Q5LAR7_BACFN	Q5lar7	bacteroides
326	32	65.3	732	2	Q64R64_BACFR	Q64r64	bacteroides
327	32	65.3	754	2	Q8BRK4_MOUSE	Q8brk4	mus musculu
328	32	65.3	759	2	Q9JZ50_NEIMB	Q9jz50	neisseria m
329	32	65.3	782	2	Q74QJ8_YERPE	Q74qj8	yersinia pe
330	32	65.3	782	2	Q8CKW7_YERPE	Q8ckw7	yersinia pe
331	32	65.3	782	2	Q8ZHC8_YERPE	Q8zhc8	yersinia pe
332	32	65.3	782	2	Q8ZG45_YERPE	Q8zga5	yersinia pe
333	32	65.3	782	2	Q666K9_YERPS	Q666k9	yersinia ps
334	32	65.3	799	1	CADH8_HUMAN	P55286	homo sapien
335	32	65.3	799	1	CADH8_RAT	P97291	mus musculu
336	32	65.3	799	1	CADH8_RAT	O54800	rattus norv
337	32	65.3	799	2	Q7N251_PHOLL	Q7n251	photorhabdu
338	32	65.3	824	2	Q8FED2_ECOL6	Q8fed2	escherichia
339	32	65.3	832	2	Q5VT46_HUMAN	Q5vt46	homo sapien
340	32	65.3	832	2	Q66CB6_YERPS	Q66cb6	yersinia ps
341	32	65.3	841	2	Q7N257_PHOLL	Q7n257	photorhabdu
342	32	65.3	842	2	Q7N245_PHOLL	Q7n245	photorhabdu
343	32	65.3	854	1	DISC1_HUMAN	Q9nri5	homo sapien
344	32	65.3	857	2	Q5KX82_GEOKA	Q5kx82	geobacillus
345	32	65.3	872	2	Q5QWH4_IDILO	Q5qwh4	idiomarina
346	32	65.3	880	2	Q74IS8_LACJO	Q74is8	lactobacilli
347	32	65.3	900	2	Q66CB3_YERPS	Q66cb3	yersinia ps
348	32	65.3	910	2	Q6B390_9PROT	Q6b390	uncultured
349	32	65.3	930	2	Q8YMS8_ANASP	Q8yms8	anabaena sp
350	32	65.3	937	2	Q4U920_THEAN	Q4u920	theileria a
351	32	65.3	940	1	UVRA_VIBCH	Q9kuw5	vibrio chol
352	32	65.3	940	1	UVRA_VIBPA	Q87la0	vibrio para
353	32	65.3	952	2	Q8ZVJ8_PYRAE	Q8zvj8	pyrobaculum
354	32	65.3	969	2	Q5CJ01_CRYHO	Q5cj01	cryptospori
355	32	65.3	986	2	Q900J9_9RETR	Q900j9	ovine lenti
356	32	65.3	1045	2	P90884_CAEEL	P90884	caenorhabdi
357	32	65.3	1050	2	Q7Q1Y1_ANOGA	Q7qly1	anopheles g
358	32	65.3	1057	2	Q61DQ1_CAEER	Q6ldq1	caenorhabdi
359	32	65.3	1068	2	Q6FNM9_CANGA	Q6fnm9	candida gla
360	32	65.3	1136	2	Q5AT35_EMENI	Q5at35	aspergillus
361	32	65.3	1167	2	Q7PM00_ANOGA	Q7pm00	anopheles g
362	32	65.3	1187	2	Q7RWR6_NEUCR	Q7rwr6	neurospora
363	32	65.3	1246	1	JARD2_HUMAN	Q92833	homo sapien
364	32	65.3	1267	2	Q4NGY1_THEPA	Q4n6y1	theileria p
365	32	65.3	1571	2	Q69ZR2_MOUSE	Q69zr2	mus musculu
366	32	65.3	1620	1	HECD1_HUMAN	Q9ult8	homo sapien
367	32	65.3	1899	2	Q4QEU6_LEIMA	Q4geu6	leishmania
368	32	65.3	2148	2	Q6QTA9_9VIRU	Q6qta9	puumala vir
369	32	65.3	2156	1	RRPL_PUTMH	P27176	puumala vir
370	32	65.3	2271	2	Q4I878_GIBZE	Q4i878	gibberella
371	32	65.3	2612	2	Q86VJ1_HUMAN	Q86vj1	homo sapien
372	32	65.3	2771	2	Q4Y148_PLACH	Q4y148	plasmodium
373	31	63.3	56	2	Q95J69_PANTR	Q95j69	pan troglod
374	31	63.3	60	2	Q8FI77_ECOL6	Q8fi77	escherichia
375	31	63.3	63	2	P72432_STREQ	P72432	streptococ
376	31	63.3	88	2	Q6VSC4_ORYSA	Q6vsc4	oryza sativ
377	31	63.3	91	1	RL34_PYRAB	Q9uzj7	pyrococcus
378	31	63.3	91	1	RL34_PYRHO	O74006	pyrococcus
379	31	63.3	95	2	Q4FLQ6_9RICK	Q4flq6	candidatus
380	31	63.3	96	1	Y3713_CLOAB	P34159	clostridium
381	31	63.3	96	2	Q56BT3_9CAUD	Q56bt3	enterobacte
382	31	63.3	102	2	Q6YSC3_ORYSA	Q6ysc3	oryza sativ
383	31	63.3	110	2	Q6MBX6_PARUM	Q6mbx6	parachlamyd
384	31	63.3	113	2	Q6HNE2_BACHK	Q6hne2	bacillus th
385	31	63.3	113	2	Q81V31_BACAN	Q81v31	bacillus an
386	31	63.3	116	2	Q63FX4_BACCZ	Q63fx4	bacillus ce
387	31	63.3	117	2	Q8DSI8_STRMU	Q8dsi8	streptococ
388	31	63.3	118	1	RS13_SHEON	Q8ek48	shewanella
389	31	63.3	120	1	RS13_NEIMA	P66385	neisseria m
390	31	63.3	120	2	Q5F5U9_NEIG1	P66385	neisseria m
391	31	63.3	120	1	Q7PAI3_RICSI	Q5f5u9	neisseria g
392	31	63.3	127	2	Q7ZE94_RICPR	Q7pai3	rickettsia
393	31	63.3	127	2	HVA22_HORVU	Q9ze94	rickettsia
394	31	63.3	130	1	Q7XAP6_WHEAT	Q07764	hordeum vul
395	31	63.3	130	2	Q68XS6_RICTY	Q7xap6	triticum ae
396	31	63.3	131	2	Q68XS6_RICTY	Q68xs6	rickettsia

397 31 63.3 131 2 Q90276_9TELE Q90276 boulergerel 470 31 63.3 274 2 Q4MLJ8_BACCE Q4mlj8 bacillus ce
398 31 63.3 133 2 Q774U3_9GEMI Q774u3 cotton leaf 471 31 63.3 274 2 Q6HH25_BACHK Q6hh25 bacillus th
399 31 63.3 133 2 Q774U7_9GEMI Q774u7 cotton leaf 472 31 63.3 274 2 Q735Q4_BACC1 Q735q4 bacillus ce
400 31 63.3 133 2 Q77M72_9GEMI Q77m72 cotton leaf 473 31 63.3 274 2 Q813F6_BACCR Q813f6 bacillus ce
401 31 63.3 133 2 Q8UYG4_9GEMI Q8uyg4 cotton leaf 474 31 63.3 274 2 Q81NV3_BACAN Q81nv3 bacillus an
402 31 63.3 133 2 Q8V0H7_9GEMI Q8v0h7 cotton leaf 475 31 63.3 274 2 Q639P3_BACCZ Q639p3 bacillus ce
403 31 63.3 133 2 Q9IN50_9GEMI Q9in50 cotton leaf 476 31 63.3 275 2 Q9M5A7_MAIZE Q9m5a7 zea mays (m
404 31 63.3 134 2 Q5L4H1_9GEMI Q5l4h1 east africa 477 31 63.3 282 2 Q6AQ8 DESPS Q6aq8 desulfotale
405 31 63.3 136 2 Q660K8_BORGA Q660k8 borrelia ga 478 31 63.3 283 2 Q9LQ05_ARATH Q9lq05 arabidopsis
406 31 63.3 139 2 Q92JJ0_RICCN Q92jj0 rickettsia 479 31 63.3 283 2 Q8LG94_ARATH Q8lg94 arabidopsis
407 31 63.3 141 2 Q592R8_LYMST Q592r8 lymnaea sta 480 31 63.3 284 1 Y1834_THETN Y1834 thetn
408 31 63.3 149 2 Q485T6_TETNG Q485t6 tetraodon n 481 31 63.3 284 2 Q8GX85_ARATH Q8gxb5 arabidopsis
409 31 63.3 150 2 Q8RFE1_FUSNN Q8rfel fusobacteri 482 31 63.3 291 2 Q9UZI7_PYRAB Q9uzi7 pyrococcus
410 31 63.3 151 2 Q750I1_ASHGO Q750i1 ashbya goss 483 31 63.3 291 2 Q9FPQ8_ARATH Q9fpq8 arabidopsis
411 31 63.3 153 2 Q8ERJ5_OCEIH Q8erj5 oceanobacil 484 31 63.3 292 2 Q97V56_SULSO Q97v56 sulfolobus
412 31 63.3 158 1 GRDA2_EUBAC Q325l8 eubacterium 485 31 63.3 292 2 Q8Y536_LISMO Q8y536 listeria mo
413 31 63.3 164 2 Q9L770_STREQ Q9l770 streptococc 486 31 63.3 292 2 Q71XC5_LISMF Q71xc5 listeria mo
414 31 63.3 165 2 Q8E286_STRA5 Q8e286 streptococc 487 31 63.3 293 2 Q8TL92_METAC Q8tl92 methanosarc
415 31 63.3 165 2 Q8E7P4_STRA3 Q8e7p4 streptococc 488 31 63.3 294 1 HD1_BRANA HD1 brana
416 31 63.3 165 2 Q5X4V8_LEGPA Q5x4v8 legionella 489 31 63.3 295 1 ST1E1_BOVIN ST1E1 bovin
417 31 63.3 166 2 Q9I2D7_PSEAE Q9i2d7 pseudomonas 490 31 63.3 300 2 Q7ME45_VIBVY Q7me45 vibrio vuln
418 31 63.3 174 2 Q720C0_LISMF Q720c0 listeria mo 491 31 63.3 301 2 Q4SF76_TETNG Q4sf76 tetraodon n
419 31 63.3 175 2 Q5ECC7_9PARA Q5ecc7 sendai viru 492 31 63.3 311 1 REPD_STAUA REPD staa
420 31 63.3 175 2 Q5ECD4_9PARA Q5ecc4 sendai viru 493 31 63.3 314 1 ALDR_HUMAN ALDR human
421 31 63.3 176 2 Q9LAV3_9STRE Q9lav3 streptococc 494 31 63.3 315 1 Q6ICP2_HUMAN Q6icp2 homo sapien
422 31 63.3 181 2 Q5ECC8_9PARA Q5ecc8 sendai viru 495 31 63.3 316 2 Q6FGA4_HUMAN Q6fga4 homo sapien
423 31 63.3 181 2 Q74HY5_LACJO Q74hy5 lactobacill 496 31 63.3 316 2 Q5U031_HUMAN Q5u031 homo sapien
424 31 63.3 187 2 Q4SI71_ARATH Q4si71 arabidopsis 497 31 63.3 316 2 Q5RAB3_PONPY Q5rab3 pongo pygma
425 31 63.3 188 2 Q5HX51_CAMJR Q5hx51 campylobact 498 31 63.3 316 2 Q8THQ3_METAC Q8thq3 methanosarc
426 31 63.3 189 2 Q9PJ12_CAMJE Q9pj12 campylobact 499 31 63.3 318 2 Q4QJD6_LEIMA Q4qjd6 leishmania
427 31 63.3 189 2 Q7R7T9_PLAYO Q7r7t9 plasmodium 500 31 63.3 323 2 Q5X2H8_LEGPA Q5x2h8 legionella
428 31 63.3 190 2 Q6F265_MESFL Q6f265 mesoplasma 501 31 63.3 326 2 Q6NRJ5_XENLA Q6nrj5 xenopus lae
429 31 63.3 199 2 Q8Q7V3_9HIV1 Q8q7v3 human immun 502 31 63.3 330 2 Q6SI33_9BACT Q6si33 uncultured
430 31 63.3 200 2 Q5ECC9_9PARA Q5ecc9 sendai viru 503 31 63.3 330 2 Q8CZA7_STRR6 Q8cza7 streptococc
431 31 63.3 204 2 Q5ECD6_9PARA Q5ecd6 sendai viru 504 31 63.3 334 2 Q6BM70_DEBHA Q6bm70 debaryomyce
432 31 63.3 204 2 Q8CY81_STRR6 Q8cy81 streptococc 505 31 63.3 334 2 Q81MY1_BACAN Q81my1 bacillus an
433 31 63.3 206 2 Q97NG6_STRPN Q97ng6 streptococc 506 31 63.3 335 2 Q9Z746_CHLPN Q9z746 chlamydia p
434 31 63.3 206 2 Q7ZWN7_XENLA Q7zwn7 xenopus lae 507 31 63.3 341 2 Q66J96_XENLA Q66j96 xenopus lae
435 31 63.3 206 2 Q8R772_THETN Q8r772 thermoanaer 508 31 63.3 349 2 Q66JL0_XENTR Q66jl0 xenopus tro
436 31 63.3 209 2 Q9EUP5_THETH Q9eup5 thermus the 509 31 63.3 349 2 Q5H6B4_XANOR Q5h6b4 xanthomonas
437 31 63.3 211 2 Q5SKT0_THET8 Q5skt0 thermus the 510 31 63.3 350 2 Q4UNW3_XANCP Q4unw3 xanthomonas
438 31 63.3 211 2 Q72L68_THET2 Q72l68 thermus the 511 31 63.3 350 2 Q8P3E1_XANCP Q8p3e1 xanthomonas
439 31 63.3 212 2 Q8T9Z4_DROOR Q8t9z4 drosophila 512 31 63.3 352 2 Q9IU81_9HIV1 Q9iu81 human immun
440 31 63.3 212 2 Q73J71_TREDE Q73j71 treponema d 513 31 63.3 352 2 Q5I5X7_ENTHI Q5i5x7 entamoeba h
441 31 63.3 213 1 KTHY_MYCGA K73ad7 mycoplasma 514 31 63.3 354 2 Q5GFV9_9HIV1 Q5gfv9 human immun
442 31 63.3 214 1 C_SEND6 P14253 sendai viru 515 31 63.3 361 2 Q50MW3_ENTHI Q50mw3 entamoeba h
443 31 63.3 214 1 C_SENDA Q9due0 sendai viru 516 31 63.3 362 2 Q9D511_MOUSE Q9d511 mus musculu
444 31 63.3 214 1 C_SENDF P14254 sendai viru 517 31 63.3 363 2 Q9CWV4_MOUSE Q9cwv4 mus musculu
445 31 63.3 214 1 C_SENDH P04861 sendai viru 518 31 63.3 363 2 Q61WV9_CAEBR Q61wv9 caenorhabdi
446 31 63.3 214 1 C_SENDO P69738 sendai viru 519 31 63.3 366 2 Q9F763_BACFR Q9f763 bacteroides
447 31 63.3 214 1 C_SENDO O55527 sendai viru 520 31 63.3 366 2 Q5LE55_BACFN Q5le55 bacteroides
448 31 63.3 214 1 C_SENDZ P04862 sendai viru 521 31 63.3 366 2 Q64V97_BACFR Q64v97 bacteroides
449 31 63.3 215 2 Q5ECD3_9PARA Q5ecd3 sendai viru 522 31 63.3 369 2 Q5HZE4_RAT Q5hze4 rattus norv
450 31 63.3 215 2 Q5ECE0_9PARA Q5ece0 sendai viru 523 31 63.3 369 2 Q8BXF6_MOUSE Q8bxf6 mus musculu
451 31 63.3 225 2 Q7YSV7_DROSI Q7ysv7 drosophila 524 31 63.3 369 2 Q9CQT1_MOUSE Q9cqt1 m mus muscu
452 31 63.3 225 2 Q7YSW6_DROSI Q7ysw6 drosophila 525 31 63.3 374 2 Q75IY0_ORYSA Q75iy0 oryza sativ
453 31 63.3 225 2 Q7YUD0_DROSI Q7yud0 drosophila 526 31 63.3 374 2 Q50LM9_ENTHI Q50lm9 entamoeba h
454 31 63.3 225 2 Q7YUD1_DROSI Q7yud1 drosophila 527 31 63.3 378 2 Q8GXU3_ARATH Q8gxu3 arabidopsis
455 31 63.3 225 2 Q7YUD2_DROSI Q7yud2 drosophila 528 31 63.3 378 2 Q581Q3_9TRYP Q581q3 trypanosoma
456 31 63.3 225 2 Q7YUD3_DROSI Q7yud3 drosophila 529 31 63.3 380 2 Q5METK_AMOPS Q5metk amoeba prot
457 31 63.3 225 2 Q9BJZ7_DROSE Q9bjz7 drosophila 530 31 63.3 381 1 Q84L19_ORYSA Q84l19 oryza sativ
458 31 63.3 225 2 Q9BJZ8_DROMA Q9bjz8 drosophila 531 31 63.3 384 2 Q6L8F7_ORYSA Q6l8f7 oryza sativ
459 31 63.3 225 2 Q9BJZ9_DROSI Q9bjz9 drosophila 532 31 63.3 384 2 Q4MTU8_BACCE Q4mtu8 bacillus ce
460 31 63.3 225 2 Q50VX7_ENTHI Q50vx7 entamoeba h 533 31 63.3 384 2 Q81DW8_BACCR Q8ldw8 bacillus ce
461 31 63.3 231 2 Q8VY70_ARATH Q8vy70 arabidopsis 534 31 63.3 385 2 Q5JLE5_ORYSA Q5jle5 oryza sativ
462 31 63.3 237 2 Q8T9Y9_9DIPT Q8t9y9 drosophila 535 31 63.3 385 2 Q7D6P7_MYCTU Q7d6p7 mycobacteri
463 31 63.3 238 2 Q8SQS5_ENCCU Q8sqes5 encephalito 536 31 63.3 386 2 Q33229_MYCTU Q33229 mycobacteri
464 31 63.3 239 2 Q8T9Z0_DROTK Q8t9z0 drosophila 537 31 63.3 386 2 Q7TY12_MYCBO Q7ty12 mycobacteri
465 31 63.3 241 2 Q8PTA1_METMA Q8ptal methanosarc 538 31 63.3 389 2 O13025_XENLA O13025 xenopus lae
466 31 63.3 248 2 Q59EL5_HUMAN Q59el5 galdieria s 539 31 63.3 391 2 Q8EW51_MYCPE Q8ew51 mycoplasma
467 31 63.3 259 1 Q9KMK5_VIBCH Q9kmk5 vibrio chol 540 31 63.3 391 2 Q5X3J7_LEGPA Q5x3j7 legionella
468 31 63.3 260 2 Q9KMK5_VIBCH Q9kmk5 vibrio chol 541 31 63.3 391 2 Q4ZRX3_PSESY Q4zrx3 pseudomonas
469 31 63.3 268 2 Q9KMK5_VIBCH Q9kmk5 vibrio chol 542 31 63.3 392 2 Q4ZRX3_PSESY Q4zrx3 pseudomonas

543	31	63.3	396	1	IMPK_MOUSE	Q7ttl6	mus musculus	616	31	63.3	642	2	Q528P6_MAGGR	Q528p6	magnaporthe
544	31	63.3	396	1	IMPK_RAT	Q99ni4	rattus norv	617	31	63.3	674	2	Q4Z605_PLABE	Q4z605	plasmodium
545	31	63.3	399	2	Q5ZT03_LEGPH	Q5zt03	legionella	618	31	63.3	675	2	Q6YZ90_ORYSA	Q6yz90	oryza sativ
546	31	63.3	402	2	Q66J00_XENLA	Q66j00	xenopus lae	619	31	63.3	678	2	Q6GNS6_XENLA	Q6gns6	xenopus lae
547	31	63.3	411	2	Q7PZU6_ANOGA	Q7pzu6	anopheles g	620	31	63.3	686	2	Q4WJ16_ASPFU	Q4wj16	aspergillus
548	31	63.3	413	2	Q5V6Q7_HALMA	Q5v6q7	haloarcula	621	31	63.3	698	2	Q7RDI7_PLAYO	Q7rdi7	plasmodium
549	31	63.3	415	2	Q8ZCK7_YERPE	Q8zck7	yersinia pe	622	31	63.3	716	2	Q5KZY8_GEOKA	Q5kzy8	geobacillus
550	31	63.3	415	2	Q668N8_YERPS	Q668n8	yersinia ps	623	31	63.3	716	2	Q4SKI9_TETNG	Q4ski9	tetraodon n
551	31	63.3	416	1	IMPK_HUMAN	Q8nfu5	homo sapien	624	31	63.3	718	2	Q4I570_GIBZE	Q4i570	gibberella
552	31	63.3	426	2	Q55F21_DICDI	Q55f21	dictyosteli	625	31	63.3	725	2	Q4I347_TNVD	Q4i347	tobacco nec
553	31	63.3	427	1	TCO2_RAT	Q9r0d6	rattus norv	626	31	63.3	726	2	Q60Z15_CAEBR	Q60z15	caenorhabdi
554	31	63.3	430	1	TCO2_MOUSE	O88968	mus musculus	627	31	63.3	756	2	Q55D97_DICDI	Q55d97	dictyosteli
555	31	63.3	430	1	YAE9_SCHPO	Q09849	schizosacch	628	31	63.3	758	2	Q6CFP0_YARLI	Q6cfp0	yarrowia li
556	31	63.3	430	2	Q5SQ21_MOUSE	Q5sq21	mus musculus	629	31	63.3	764	2	Q4UHP6_THEAN	Q4uhp6	theileria a
557	31	63.3	431	2	Q66I16_XENTR	Q66i16	xenopus tro	630	31	63.3	773	2	Q4MS75_BACCE	Q4ms75	bacillus ce
558	31	63.3	433	2	Q74MI6_NANEQ	Q74mi6	nanoarchaeu	631	31	63.3	773	2	Q72ZV6_BACC1	Q72zv6	bacillus ce
559	31	63.3	434	2	Q84VM2_XENLA	Q84vm2	xenopus lae	632	31	63.3	773	2	Q81LC1_BACAN	Q81lc1	bacillus an
560	31	63.3	435	2	Q8I5V3_PLAF7	Q8i5v3	plasmodium	633	31	63.3	774	2	Q65GJ6_BACLD	Q65gj6	bacillus li
561	31	63.3	437	2	Q5X7E3_LEGPA	Q5x7e3	legionella	634	31	63.3	775	2	Q58ME7_9CAUD	Q58me7	cyanophage
562	31	63.3	437	2	Q5ZXX0_LEGPH	Q5zxx0	legionella	635	31	63.3	776	2	Q5DMK3_BPT5	Q5dmk3	bacterioph
563	31	63.3	450	2	Q96MS1_HUMAN	Q96ms1	homo sapien	636	31	63.3	776	2	Q6HSS4_BACAN	Q6hss4	bacillus an
564	31	63.3	451	1	Y4I3_STRAS	Q8ele4	streptococc	637	31	63.3	776	2	Q817Q4_BACCR	Q817q4	bacillus ce
565	31	63.3	451	1	Y448_STRAS	Q8e6w1	streptococc	638	31	63.3	776	2	Q6HD56_BACHK	Q6hd56	bacillus th
566	31	63.3	452	2	Q75IY1_ORYSA	Q75iy1	oryza sativ	639	31	63.3	776	2	Q633X4_BACCC	Q633x4	bacillus ce
567	31	63.3	452	2	Q56630_VIBCH	Q56630	vibrio chol	640	31	63.3	779	2	Q55GX4_DICDI	Q55gx4	dictyosteli
568	31	63.3	452	2	Q6P846_XENTR	Q6p846	xenopus tro	641	31	63.3	782	2	Q789F1_CHICK	Q789f1	gallus gall
569	31	63.3	459	2	Q8EQ57_OCEIH	Q8eq57	oceanobacil	642	31	63.3	798	2	Q4S3L5_TETNG	Q4s3l5	tetraodon n
570	31	63.3	459	2	Q5AWT5_EMENI	Q5awt5	aspergillus	643	31	63.3	806	2	Q751P7_ASHGO	Q751p7	ashbya goss
571	31	63.3	460	2	Q54Q24_DICDI	Q54q24	dictyosteli	644	31	63.3	811	2	Q6QGJ1_BPT5	Q6qgj1	bacterioph
572	31	63.3	462	2	Q4MQQ8_BACCE	Q4mqq8	anaeromyxob	645	31	63.3	811	2	Q66LM9_BPT5	Q66lw9	bacterioph
573	31	63.3	464	2	Q4NT34_9DELT	Q4nt34	dictyosteli	646	31	63.3	833	2	Q5KE62_CRYNE	Q5ke62	cryptococcu
574	31	63.3	466	2	Q4I5V1_GIBZE	Q4i5v1	gibberella	647	31	63.3	842	2	Q50ML6_ENTHI	Q50ml6	entamoeba h
575	31	63.3	474	2	Q5P0X2_AZOSE	Q5p0x2	azoarcus sp	648	31	63.3	848	2	Q91PK3_WSSV	Q91pk3	white spot
576	31	63.3	477	2	Q7NYV2_CHRVO	Q7nyv2	chromobacte	649	31	63.3	848	2	Q9QAL9_WSSV	Q9qal9	white spot
577	31	63.3	477	2	Q6BTT2_DEBHA	Q6btt2	debaromyce	650	31	63.3	852	2	Q50VA1_ENTHI	Q50va1	entamoeba h
578	31	63.3	479	2	Q5JLGO_ORYSA	Q5jlg0	oryza sativ	651	31	63.3	856	1	FBIC_MYCTU	Q7u0g9	mycobacteri
579	31	63.3	486	2	Q6JLJ6_ORYSA	Q6jlg0	oryza sativ	652	31	63.3	856	1	FBIC_MYCTU	O50429	mycobacteri
580	31	63.3	488	1	ARLY2_RHIME	Q92vm6	rhizobium m	653	31	63.3	860	2	Q42351_CHICK	Q42351	gallus gall
581	31	63.3	488	1	U2AF2_CAEBR	P90727	caenorhabdi	654	31	63.3	861	2	Q50MR0_ENTHI	Q50mr0	entamoeba h
582	31	63.3	496	2	Q5GCA4_RHOGE	Q5gca4	rhodocyclus	655	31	63.3	866	2	Q50PN8_ENTHI	Q50pn8	entamoeba h
583	31	63.3	496	2	Q5WVS0_LEGPL	Q5wvs0	legionella	656	31	63.3	866	2	Q510I5_ENTHI	Q510i5	entamoeba h
584	31	63.3	496	2	Q5X4D5_LEGPA	Q5x4d5	legionella	657	31	63.3	866	2	Q513Q9_ENTHI	Q513q9	entamoeba h
585	31	63.3	496	2	Q5ZUL6_LEGPH	Q5zul6	legionella	658	31	63.3	872	2	Q4XPH4_PLACH	Q4xph4	plasmodium
586	31	63.3	501	2	Q6CJD5_KLULA	Q6cjd5	kluyveromyc	659	31	63.3	897	2	Q9V329_DROME	Q9v329	drosophila
587	31	63.3	501	2	Q8G7L4_BIFLO	Q8g7l4	bifidobacte	660	31	63.3	920	2	Q6ER88_ORYSA	Q6er88	oryza sativ
588	31	63.3	501	2	Q896U3_CLOTE	Q896u3	clostridium	661	31	63.3	924	2	Q9CHC9_LACLA	Q9chc9	lactococcus
589	31	63.3	502	2	Q9VFS1_DROME	Q9vfs1	drosophila	662	31	63.3	928	2	Q660E1_BORGA	Q660e1	borrelia ga
590	31	63.3	507	2	Q9N5Z9_CABEL	Q9n5z9	caenorhabdi	663	31	63.3	929	2	Q7R7Q2_PLAYO	Q7r7q2	plasmodium
591	31	63.3	514	2	Q9LJY6_ARATH	Q9ljy6	arabidopsis	664	31	63.3	932	2	O51674_BORBU	O51674	borrelia bu
592	31	63.3	515	2	Q4LX93_9BURK	Q4lx93	burkholderi	665	31	63.3	946	1	UVRA_ZYMMO	O31151	zymomonas m
593	31	63.3	518	2	Q9YKG9_XENLA	Q9ykg9	xenopus lae	666	31	63.3	954	2	Q83EP1_COXBU	Q83ep1	coxiella bu
594	31	63.3	520	2	Q4LDN2_XENLA	Q4ldn2	xenopus lae	667	31	63.3	955	2	O8IYE0_HUMAN	O8iye0	homo sapien
595	31	63.3	527	2	Q87G39_VIBPA	Q87g39	vibrio para	668	31	63.3	961	2	Q9JGP1_9VIRU	Q9jgp1	apple laten
596	31	63.3	535	2	Q4Z6E2_PLABE	Q4z6e2	plasmodium	669	31	63.3	974	1	GLNE_BARHE	Q6g488	bartonella
597	31	63.3	537	2	Q804W8_FUGRU	Q804w8	fugu rubrip	670	31	63.3	974	2	Q66H60_RAT	Q66h60	rattus norv
598	31	63.3	543	2	Q95JS5_MACFA	Q95js5	macaca fasc	671	31	63.3	987	2	Q60D46_SOLDE	Q60d46	solanum dem
599	31	63.3	549	2	Q67M51_SYMTH	Q67m51	symbiobacte	672	31	63.3	998	2	Q50ZM8_ENTHI	Q50zm8	entamoeba h
600	31	63.3	557	2	Q54AN6_DICDI	Q54an6	dictyosteli	673	31	63.3	1012	2	Q559G8_DICDI	Q559g8	dictyosteli
601	31	63.3	557	2	Q73766_XENLA	Q73766	xenopus lae	674	31	63.3	1016	2	Q6ZTG8_HUMAN	Q6ztg8	homo sapien
602	31	63.3	564	2	Q4I6V9_GIBZE	Q4i6v9	gibberella	675	31	63.3	1020	2	Q66J08_XENLA	Q66j08	xenopus lae
603	31	63.3	578	2	Q9RGC5_MYCHY	Q9rgc5	mycoplasma	676	31	63.3	1043	2	O8IYYP4_ANASP	O8iyypp4	anabaena sp
604	31	63.3	590	2	Q9KVA7_VIBCH	Q9kva7	vibrio chol	677	31	63.3	1061	2	Q8I5M3_PLAF7	Q8i5m3	plasmodium
605	31	63.3	593	2	Q8YPL5_ANASP	Q8ypl5	anabaena sp	678	31	63.3	1074	2	Q5XH56_XENLA	Q5xh56	xenopus lae
606	31	63.3	597	2	Q6FXJ1_CANGA	Q6fxj1	candida gla	679	31	63.3	1076	1	YE38_SCHPO	O13892	schizosacch
607	31	63.3	600	2	Q4SAX3_TETNG	Q4sax3	tetraodon n	680	31	63.3	1077	2	Q7KQ94_DICDI	Q7kq94	dictyosteli
608	31	63.3	601	2	Q4YB39_PLABE	Q4yb39	plasmodium	681	31	63.3	1098	2	Q6CPF5_KLULA	Q6cpf5	kluyveromyc
609	31	63.3	603	2	Q7RHJ1_PLAYO	Q7rhj1	plasmodium	682	31	63.3	1116	2	O15725_DICDI	O15725	dictyosteli
610	31	63.3	617	2	Q6LKL8_PHOPR	Q6lkl8	photobacter	683	31	63.3	1153	2	Q5SH04_DICDI	Q5sh04	dictyosteli
611	31	63.3	626	2	Q503P8_BRARE	Q503p8	brachydanio	684	31	63.3	1170	2	Q8IP22_DROME	Q8ip22	drosophila
612	31	63.3	627	2	Q7T2M4_CARAU	Q7t2m4	carassius a	685	31	63.3	1174	2	Q932X6_LACLA	Q932x6	lactococcus
613	31	63.3	630	2	Q4I5G9_GIBZE	Q4i5g9	gibberella	686	31	63.3	1174	2	Q7BS52_9LACT	Q7bb52	lactococcus
614	31	63.3	632	1	ASNB_BACSU	P54420	bacillus su	687	31	63.3	1176	2	Q8T2S7_DICDI	Q8t2s7	dictyosteli
615	31	63.3	636	2	Q5TW56_ANOGA	Q5tw56	anopheles g	688	31	63.3	1203	2	Q5FC45_CABEL	O5fc45	caenorhabdi

689	31	63.3	1205	2	P92021_CAEEL	P92021_caenorhabdi	762	30	61.2	151	2	Q6C242_YARLI	Q6c242 yarrowia li
690	31	63.3	1210	2	Q8EQJ8_OCEIH	Q8eqj8 oceanobacil	763	30	61.2	151	2	Q7VI02_HELHP	Q7vi02 helicobacte
691	31	63.3	1215	2	Q5BIB2_DROME	Q5bib2 drosophila	764	30	61.2	151	2	Q65NH9_BACLD	Q65nh9 bacillus li
692	31	63.3	1257	2	Q5SP94_CRYNE	Q5sp94 cryptococcu	765	30	61.2	154	2	Q26394_METTH	Q26394 methanobact
693	31	63.3	1280	2	Q6FNG0_CANGA	Q6fng0 candida gla	766	30	61.2	154	2	Q7NJZ4_GLOVI	Q7njz4 gloebobacter
694	31	63.3	1307	2	Q4QDX6_LEIMA	Q4qdx6 leishmania	767	30	61.2	162	2	Q72CA7_DESVH	Q72ca7 desulfovibr
695	31	63.3	1317	2	Q7RH57_PLAYO	Q7rh57 plasmodium	768	30	61.2	172	2	Q6A5D4_PROAC	Q6a5d4 propionibac
696	31	63.3	1322	2	Q54TV7_DICDI	Q54tv7 dictyosteli	769	30	61.2	172	2	Q91057_9TELE	Q91057 hemiodus sp
697	31	63.3	1322	2	Q55D46_DICDI	Q55d46 dictyosteli	770	30	61.2	175	2	Q4MRF1_BACEE	Q4mrf1 bacillus ce
698	31	63.3	1322	2	Q55GX9_DICDI	Q55gx9 dictyosteli	771	30	61.2	175	2	Q6HJV9_BACHK	Q6hvj9 bacillus th
699	31	63.3	1335	2	Q55GX6_DICDI	Q55gx6 dictyosteli	772	30	61.2	175	2	Q81RN5_BACAN	Q81rn5 bacillus an
700	31	63.3	1382	2	Q4SM61_TETNG	Q4sm61 tetraodon n	773	30	61.2	175	2	Q91140_nannobrycon	Q91140 nannobrycon
701	31	63.3	1423	2	Q5T2J2_HUMAN	Q5t2j2 homo sapien	774	30	61.2	178	2	Q4JA49_SULAC	Q4ja49 sulfolobus
702	31	63.3	1461	2	Q7RBI1_PLAYO	Q7rb11 plasmodium	775	30	61.2	178	2	Q91083_9TELE	Q91083 leporinus s
703	31	63.3	1509	2	Q8S6P1_ORYSA	Q8s6p1 oryza sativ	776	30	61.2	181	2	Q90394_9TELE	Q90394 chalceus sp
704	31	63.3	1534	2	Q4TBT5_TETNG	Q4tbt5 tetraodon n	777	30	61.2	181	2	Q91109_9TELE	Q91109 metynnis sp
705	31	63.3	1535	2	Q5CIU2_CRYHO	Q5ciu2 cryptospori	778	30	61.2	188	2	Q5XG31_XENLA	Q5xg31 xenopus lae
706	31	63.3	1692	2	Q8IJD6_PLAF7	Q8ijd6 plasmodium	779	30	61.2	190	2	Q5SPA1_BRARE	Q5spa1 brachydanio
707	31	63.3	2057	2	Q882M4_PSESM	Q882m4 pseudomonas	780	30	61.2	192	2	Q51CP4_ENTHI	Q51cp4 entamoeba h
708	31	63.3	2087	2	Q4S488_TETNG	Q4s488 tetraodon n	781	30	61.2	194	2	Q5TSS0_HUMAN	Q5tss0 homo sapien
709	31	63.3	2190	2	Q9VJM0_DROME	Q9vjm0 drosophila	782	30	61.2	194	2	Q9KPR4_VIBCH	Q9kpr4 vibrio chol
710	31	63.3	2223	1	CAC1E_DISOM	P56699 discopyge o	783	30	61.2	194	2	Q4SS57_TETNG	Q4ss57 tetraodon n
711	31	63.3	2314	2	Q7PQ1I_ANOGA	Q7pq1i anopheles g	784	30	61.2	199	2	Q9RMF1_ZYMGO	Q9rmf1 zymomonas m
712	31	63.3	2576	2	Q8JFV5_BRARE	Q8jfv5 brachydanio	785	30	61.2	200	2	Q7M7V3_WOLSU	Q7m7v3 wolinnella s
713	30	61.2	35	2	Q8RIW2_CLODI	Q8riw2 clostridium	786	30	61.2	201	2	Q8FJG8_ECOL6	Q8fjg8 escherichia
714	30	61.2	58	2	Q7MBJ5_VIBVY	Q7mbj5 vibrio vuln	787	30	61.2	201	2	Q8XGS5_SALTI	Q8xgs5 salmonella
715	30	61.2	72	2	Q9N3J5_CAEEL	Q9n3j5 caenorhabdi	788	30	61.2	201	2	Q8R5J1_MARMO	Q8r5j1 marmota mon
716	30	61.2	79	2	Q500F1_PSESY	Q500f1 pseudomonas	789	30	61.2	205	2	Q87GF6_VIBPA	Q87gf6 vibrio para
717	30	61.2	87	2	Q05325_PSEFL	O05325 pseudomonas	790	30	61.2	207	2	Q736I2_BACC1	Q736i2 bacillus ce
718	30	61.2	89	2	Q4LHK1_9BURK	Q4lhk1 burkholderi	791	30	61.2	207	2	Q813H0_BACCR	Q813h0 bacillus ce
719	30	61.2	90	2	Q86WH1_HUMAN	Q86wh1 homo sapien	792	30	61.2	210	2	Q625G0_CABBR	Q625g0 caenorhabdi
720	30	61.2	93	2	Q30682_RHOER	Q30682 rhodococcus	793	30	61.2	210	2	Q9SYD8_ARATH	Q9ayd8 arabidopsis
721	30	61.2	93	2	Q5WYZ7_LEGPL	Q5wyz7 legionella	794	30	61.2	210	2	Q5PSE0_9CHLO	Q5pse0 halimeda in
722	30	61.2	93	2	Q5X7J9_LEGPA	Q5x7j9 legionella	795	30	61.2	213	1	STAR5_MOUSE	Q9epq7 mus musculu
723	30	61.2	93	2	Q5ZY31_LEGPH	Q5zy31 legionella	796	30	61.2	213	2	Q9CQY0_MOUSE	Q9cgy0 m mus muscu
724	30	61.2	98	2	Q88L82_PSEPK	Q88l82 pseudomonas	797	30	61.2	214	2	Q6FJU9_CANGA	Q6fju9 candida gla
725	30	61.2	101	2	Q8ND09_HUMAN	Q8nd09 homo sapien	798	30	61.2	216	2	Q4MPD0_BACCE	Q4mpd0 bacillus ce
726	30	61.2	110	2	P95874_SULSO	P95874 sulfolobus	799	30	61.2	217	2	Q4X744_PLACH	Q4x744 plasmodium
727	30	61.2	110	2	Q5C0D4_SCHJA	Q5c0d4 schistosoma	800	30	61.2	218	1	Y395_PASMU	Q9cnn2 pasteurella
728	30	61.2	111	2	Q5BFE5_EMENI	Q5bfe5 aspergillus	801	30	61.2	219	2	Q64CK5_9ARCH	Q64ck5 uncultured
729	30	61.2	111	2	Q9WMG9_PHV	Q9wmg9 prospect hi	802	30	61.2	219	2	Q5PSF0_9CHLO	Q5psf0 halimeda in
730	30	61.2	115	1	VG52_SHV21	Q01050 saimirine	803	30	61.2	220	2	Q50LX9_ENTHI	Q50lx9 entamoeba h
731	30	61.2	115	2	Q80BJ5_9GAMA	Q80bj5 saimirine	804	30	61.2	226	1	Y1442_YERPE	Q8zg70 versinia pe
732	30	61.2	115	2	Q80BJ9_9GAMA	Q80bj9 saimirine	805	30	61.2	226	1	Y1460_YERPS	Q66ce3 versinia ps
733	30	61.2	115	2	Q80BR6_9GAMA	Q80br6 saimirine	806	30	61.2	229	2	P94406_BACSU	P94406 bacillus su
734	30	61.2	116	2	Q6MS32_MYCMS	Q6ms32 mycoplasma	807	30	61.2	230	2	Q6P2D9_HUMAN	Q6p2d9 homo sapien
735	30	61.2	118	1	RS13_YERPE	Q8zj90 versinia pe	808	30	61.2	230	2	Q72QV2_LEPIC	Q72qv2 leptospira
736	30	61.2	118	2	Q664U3_YERPS	Q664u3 versinia ps	809	30	61.2	231	2	Q4XD89_PLACH	Q4xd89 plasmodium
737	30	61.2	119	2	Q8CX75_OCEIH	Q8cx75 oceanobacil	810	30	61.2	232	2	Q99P51_MOUSE	Q99p51 mus musculu
738	30	61.2	120	2	Q7NQH4_CHRVO	Q7nqh4 chromobacte	811	30	61.2	232	2	Q86WH2_HUMAN	Q86wh2 homo sapien
739	30	61.2	123	2	Q95Y08_CAEEL	Q95y08 caenorhabdi	812	30	61.2	238	2	Q8NPL3_CORGL	Q8npl3 corynebacte
740	30	61.2	126	1	GRP_BOVIN	Q863c3 bos taurus	813	30	61.2	241	1	PHOU_ECOLI	P0a9k7 escherichia
741	30	61.2	126	2	Q6AIH7_DESPS	Q6aih7 desulfotale	814	30	61.2	241	1	PHOU_ENTCL	Q32488 enterobacte
742	30	61.2	129	2	Q5HML8_STAEQ	Q5hml8 staphylococ	815	30	61.2	241	1	PHOU_SHIFL	P0a9k8 shigella fl
743	30	61.2	131	2	Q61TF0_CABBR	Q61tf0 caenorhabdi	816	30	61.2	241	2	Q57HY9_SALCH	Q57hy9 salmonella
744	30	61.2	133	2	Q6DHG4_BRARE	Q6dhg4 brachydanio	817	30	61.2	241	2	Q5PKV0_SALPA	Q5pkv0 salmonella
745	30	61.2	134	1	GRP_SHEEP	P47851 ovis aries	818	30	61.2	241	2	Q7CPE7_SALTY	Q7cpe7 salmonella
746	30	61.2	134	2	Q94HX5_ORYSA	Q94hx5 oryza sativ	819	30	61.2	241	2	Q8XFI0_SALTI	Q8xfi0 salmonella
747	30	61.2	134	2	Q7XCR6_ORYSA	Q7xcr6 oryza sativ	820	30	61.2	241	2	Q8XB06_ECO57	Q8xb06 escherichia
748	30	61.2	134	2	Q4FLU0_9RICK	Q4flu0 candidatus	821	30	61.2	242	2	Q8F4Z6_LEPIN	Q8f4z6 leptospira
749	30	61.2	134	2	Q6EMG6_9GEMI	Q6emg6 east africa	822	30	61.2	243	2	Q4HTN6_CAMUP	Q4htn6 campylobact
750	30	61.2	134	2	Q9WR11_9GEMI	Q9wr11 east africa	823	30	61.2	246	2	Q8S7R9_ORYSA	Q8s7r9 oryza sativ
751	30	61.2	134	2	Q58WJ1_9GEMI	Q58wj1 east africa	824	30	61.2	247	2	Q4N8W6_THEPA	Q4n8w6 theileria p
752	30	61.2	134	2	Q58WJ7_9GEMI	Q58wj7 east africa	825	30	61.2	247	2	Q8FBT8_ECOL6	Q8fbt8 escherichia
753	30	61.2	136	2	Q753T6_ASHGO	Q753t6 ashbya goss	826	30	61.2	249	2	Q4UER8_THEAN	Q4uer8 theileria a
754	30	61.2	137	2	Q91058_9TELE	Q91058 hoptlias sp.	827	30	61.2	249	2	Q9K7I0_BACHD	Q9k7i0 bacillus ha
755	30	61.2	140	2	Q83028_DEIRA	Q83028 deinococcus	828	30	61.2	249	2	Q92V08_RHIME	Q92v08 rhizobium m
756	30	61.2	144	2	Q8YJU7_ANASP	Q8yju7 anabaena sp	829	30	61.2	251	2	Q5PSE1_9CHLO	Q5pse1 halimeda in
757	30	61.2	149	2	Q85698_STRLI	Q85698 streptomyce	830	30	61.2	251	2	Q5PSF4_9CHLO	Q5psf4 halimeda in
758	30	61.2	149	2	Q7AKR5_STRCO	Q7akr5 streptomyce	831	30	61.2	252	2	Q5PSE8_9CHLO	Q5pse8 halimeda in
759	30	61.2	149	2	Q92RB8_RHIME	Q92rb8 rhizobium m	832	30	61.2	253	2	Q50P50_ENTHI	Q50p50 entamoeba h
760	30	61.2	150	1	RS13_CANMA	P33192 candida mal	833	30	61.2	256	1	Y022_THEAC	Q9hm47 thermoplasma
761	30	61.2	150	1	RS13_SCHPO	P28189 schizosacch	834	30	61.2	256	2	Q5JEG3_PYRKO	Q5jeg3 pyrococcus

835	30	61.2	256	2	Q9UYD3_PYRAB	Q9uyd3 pyrococcus	908	30	61.2	354	2	Q95Y07_CAEEL	Q95y07 caenorhabdi
836	30	61.2	256	2	O58212_PYRHO	O58212 pyrococcus	909	30	61.2	357	2	Q6CBE2_YARLI	Q6cbe2 yarrowia li
837	30	61.2	256	2	Q5L0M4_GEOKA	Q5l0m4 geobacillus	910	30	61.2	359	2	O60759_HUMAN	O60759 homo sapien
838	30	61.2	257	2	Q8U3K7_PYRFU	Q8u3k7 pyrococcus	911	30	61.2	359	2	Q8NE32_HUMAN	Q8ne32 homo sapien
839	30	61.2	257	2	Q6HIW4_BACHK	Q6hiw4 bacillus th	912	30	61.2	361	2	Q7SZ61_BRARE	Q7sz61 brachydanio
840	30	61.2	258	2	Q5PSF1_9CHLO	Q5psf1 halimeda in	913	30	61.2	362	2	Q9GTC6_CAEEL	Q9gtc6 caenorhabdi
841	30	61.2	258	2	Q5PSF6_9CHLO	Q5psf6 halimeda in	914	30	61.2	363	2	Q5KPS1_CRYNE	Q5kps1 cryptococcu
842	30	61.2	258	2	Q6MQJ6_BDEBA	Q6mqj6 bdellovibri	915	30	61.2	364	2	Q50UT8_ENTHI	Q50ut8 entamoeba h
843	30	61.2	260	2	Q5PSB9_9CHLO	Q5psb9 halimeda si	916	30	61.2	364	2	Q8VI91_RAT	Q8vi91 rattus norv
844	30	61.2	260	2	Q5PSE4_9CHLO	Q5pse4 halimeda in	917	30	61.2	365	2	Q9LHR3_ARATH	Q9lhr3 arabidopsis
845	30	61.2	261	2	Q4Y5N2_PLACH	Q4y5n2 plasmodium	918	30	61.2	367	2	Q9KLY7_VIBCH	Q9kly7 vibrio chol
846	30	61.2	262	1	VIBA_VIBCH	Q56632 vibrio chol	919	30	61.2	370	2	Q6LIU3_PHOTO	Q6liu3 photobacter
847	30	61.2	262	2	Q5PSE5_9CHLO	Q5pse5 halimeda in	920	30	61.2	371	2	Q9VKE8_DROME	Q9vke8 drosophila
848	30	61.2	263	2	Q5PSE6_9CHLO	Q5pse6 halimeda in	921	30	61.2	373	2	Q6CPK4_KLULA	Q6cpk4 kluyveromyc
849	30	61.2	263	2	Q5PSE7_9CHLO	Q5pse7 halimeda in	922	30	61.2	375	2	Q9CXS1_MOUSE	Q9cxs1 mus musculu
850	30	61.2	263	2	Q5PSE9_9CHLO	Q5pse9 halimeda in	923	30	61.2	379	2	Q738T6_BACCI	Q738t6 bacillus ce
851	30	61.2	267	2	Q9CT54_MOUSE	Q9ct54 mus musculu	924	30	61.2	379	2	Q5PPN7_RAT	Q5ppn7 rattus norv
852	30	61.2	273	2	O83507_TREPA	O83507 treponema p	925	30	61.2	380	2	Q96ER9_HUMAN	Q96er9 homo sapien
853	30	61.2	275	2	Q6BFP6_PARTE	Q6bfp6 paramecium	926	30	61.2	382	1	METK_LEGPA	Q5x3n0 legionella
854	30	61.2	275	2	Q5PSB8_9CHLO	Q5psb8 halimeda si	927	30	61.2	382	1	METK_LEGPH	Q5zty6 legionella
855	30	61.2	275	2	Q5PSC1_9CHLO	Q5psc1 halimeda si	928	30	61.2	382	1	METK_LEGPL	Q5wv18 legionella
856	30	61.2	275	2	Q5PSC2_9CHLO	Q5psc2 halimeda si	929	30	61.2	382	2	Q4N977_THEPA	Q4n977 theileria p
857	30	61.2	275	2	Q5PSC3_9CHLO	Q5psc3 halimeda si	930	30	61.2	382	2	Q64SE2_BACFR	Q64se2 bacteroides
858	30	61.2	275	2	Q5PSC5_9CHLO	Q5psc5 halimeda si	931	30	61.2	382	2	Q9A7B7_CAUCR	Q9a7b7 caulobacter
859	30	61.2	275	2	Q5PSC7_9CHLO	Q5psc7 halimeda si	932	30	61.2	382	2	Q5L3D7_GEOKA	Q5l3d7 geobacillus
860	30	61.2	279	2	Q4RGX3_MACFA	Q4rgx3 macaca fasc	933	30	61.2	385	2	Q98K34_RHILO	Q98k34 rhizobium l
861	30	61.2	284	2	Q9ZVN5_ARATH	Q9zvn5 arabidopsis	934	30	61.2	386	2	Q552U0_CRYNE	Q552u0 cryptococcu
862	30	61.2	286	2	Q6ALV3_DESPS	Q6alv3 desulfotale	935	30	61.2	387	2	P95884_SULSO	P95884 sulfolobus
863	30	61.2	286	2	Q6N863_RHOPA	Q6n863 rhodopseudo	936	30	61.2	387	2	Q7QC53_ANOGA	Q7qc53 anopheles g
864	30	61.2	288	2	Q7TSY5_9PICO	Q7tsy5 human rhino	937	30	61.2	388	2	Q91DT2_9VIRU	Q91dt2 rice tungro
865	30	61.2	292	2	Q8XJ66_CLOPE	Q8xj66 clostridium	938	30	61.2	390	2	Q96KY9_HUMAN	Q96ky9 homo sapien
866	30	61.2	295	2	Q7CME8_BACAN	Q7cme8 bacillus an	939	30	61.2	390	2	Q5H825_STRSU	Q5h825 streptococc
867	30	61.2	295	2	Q4MKX0_BACCE	Q4mkx0 bacillus ce	940	30	61.2	390	2	Q4NMY4_9DELT	Q4nmy4 anaeromyxob
868	30	61.2	295	2	Q9X364_BACAN	Q9x364 bacillus an	941	30	61.2	394	2	Q7QU72_GIALA	Q7qu72 giardia lam
869	30	61.2	297	2	Q4UGH8_THEAN	Q4ugh8 theileria a	942	30	61.2	396	2	Q9HP92_HALSA	Q9hp92 halobacteri
870	30	61.2	298	1	ERA_XANCP	Q8pb51 xanthomonas	943	30	61.2	397	2	Q4UHL9_THEAN	Q4uhl9 theileria a
871	30	61.2	298	2	Q7S3S2_NEUCR	Q7s3s2 neurospora	944	30	61.2	401	2	Q5T5S1_HUMAN	Q5t5s1 homo sapien
872	30	61.2	298	2	Q4USP8_XANCP	Q4usf8 xanthomonas	945	30	61.2	403	2	Q8ECA4_SHEON	Q8eca4 shewanella
873	30	61.2	299	1	ERA_XANAC	Q8pmu9 xanthomonas	946	30	61.2	404	2	Q5E9Q3_BOVIN	Q5e9q3 bos taurus
874	30	61.2	299	2	Q5HIR1_XANOR	Q5hlr1 xanthomonas	947	30	61.2	404	2	Q5E9W7_BOVIN	Q5e9w7 bos taurus
875	30	61.2	301	2	Q5CES2_CRYHO	Q5ce52 cryptospori	948	30	61.2	404	2	Q5EA69_BOVIN	Q5ea69 bos taurus
876	30	61.2	302	2	Q9HA01_HUMAN	Q9ha01 homo sapien	949	30	61.2	404	2	Q6NDN6_RHOPA	Q6ndn6 rhodopseudo
877	30	61.2	303	1	CBPB_ASTFL	P04069 astacus flu	950	30	61.2	406	1	DAC_STRSR	P15555 streptomyc
878	30	61.2	305	2	Q54CE1_DICDI	Q54ce1 dictyosteli	951	30	61.2	406	2	Q4QQL1_MOUSE	Q4qql1 mus musculu
879	30	61.2	305	2	Q9GYR2_CAEEL	Q9gyr2 caenorhabdi	952	30	61.2	409	2	O61103_TRYCR	O61103 trypanosoma
880	30	61.2	306	2	Q61LR8_CAEEL	Q61lr8 caenorhabdi	953	30	61.2	415	2	Q4S8M8_TETNG	Q4s8m8 tetraodon n
881	30	61.2	309	2	Q4LLB7_9BURK	Q4llb7 burkholderi	954	30	61.2	417	2	Q6YPR6_ONYPE	Q6ypr6 onion yello
882	30	61.2	313	1	MDH_BACTN	Q8a0w0 bacteroides	955	30	61.2	418	2	Q8TK44_METAC	Q8tk44 methanosarc
883	30	61.2	313	2	Q7TRP1_MOUSE	Q7trp1 mus musculu	956	30	61.2	421	2	Q70J68_STRGR	Q70j68 streptomyc
884	30	61.2	314	2	Q6JVA7_ANTPO	Q6jva7 antheraea p	957	30	61.2	422	2	Q6ZUX0_HUMAN	Q6zux0 homo sapien
885	30	61.2	315	2	Q5JG61_PYRKO	Q5jg61 pyrococcus	958	30	61.2	422	2	Q7R4K4_GIALA	Q7r4k4 giardia lam
886	30	61.2	316	2	Q50T39_ENTHI	Q50t39 entamoeba h	959	30	61.2	423	2	Q5CZ27_CRYPV	Q5cz27 cryptospori
887	30	61.2	318	2	Q7NIK9_GLOVI	Q7nik9 gloeobacter	960	30	61.2	423	2	Q5CMB9_CRYHO	Q5cmb9 cryptospori
888	30	61.2	319	2	Q4H5T7_9DEIO	Q4h5t7 deinococcus	961	30	61.2	427	2	Q7MT81_PORGI	Q7mt81 porphyromon
889	30	61.2	321	1	RLAO_CHERU	P29764 chenopodium	962	30	61.2	430	1	SYH_LACLA	Q9ce78 lactococcus
890	30	61.2	323	2	Q6BG45_PARTE	Q6bg45 paramecium	963	30	61.2	432	2	Q6G8M2_STAAS	Q6g8m2 staphylococ
891	30	61.2	324	1	MURB_RHIME	Q92nm1 rhizobium m	964	30	61.2	432	2	Q6GG02_STAAR	Q6gg02 staphylococ
892	30	61.2	324	2	Q5BBC7_EMENI	Q5bbc7 aspergillus	965	30	61.2	432	2	Q5HF69_STAAC	Q5hf69 staphylococ
893	30	61.2	325	2	Q50N53_ENTHI	Q50n53 entamoeba h	966	30	61.2	432	2	Q7A0N2_STAAN	Q7a0n2 staphylococ
894	30	61.2	326	2	Q7Q7V1_ANOGA	Q7q7v1 anopheles g	967	30	61.2	432	2	Q7A554_STAAN	Q7a554 staphylococ
895	30	61.2	326	2	Q50PQ4_ENTHI	Q50pq4 entamoeba h	968	30	61.2	432	2	Q99TF8_STAAM	Q99tf8 staphylococ
896	30	61.2	327	1	TRXB_PNEJI	Q8j0u0 pneumocysti	969	30	61.2	433	2	Q6FTS3_CANGA	Q6fts3 candida gla
897	30	61.2	331	2	Q9M4Z5_SPIOL	Q9m4z5 spinacia ol	970	30	61.2	433	2	Q50ND4_ENTHI	Q50nd4 entamoeba h
898	30	61.2	331	2	Q6DB83_ERWCT	Q6db83 erwinia car	971	30	61.2	434	2	Q9LH92_ARATH	Q9lh92 arabidopsis
899	30	61.2	334	2	Q6N255_RHOPA	Q6n255 rhodopsendo	972	30	61.2	434	2	Q9LH93_ARATH	Q9lh93 arabidopsis
900	30	61.2	335	2	Q22666_CAEEL	Q22666 caenorhabdi	973	30	61.2	440	2	Q74BH4_GEOSL	Q74bh4 geobacter s
901	30	61.2	335	2	Q7UWQ7_RHOBA	Q7uwg7 rhodopirell	974	30	61.2	445	2	Q9MBG7_ARATH	Q9mbg7 arabidopsis
902	30	61.2	335	2	Q4RSF1_TETNG	Q4rsf1 tetraodon n	975	30	61.2	446	1	Y1571_PASMU	Q9ckn9 pasteurella
903	30	61.2	336	2	Q5FHY0_LACAC	Q5fhy0 lactobacill	976	30	61.2	448	2	Q95TK5_DROME	Q95tk5 drosophila
904	30	61.2	339	2	Q8TF18_HUMAN	Q8tf18 homo sapien	977	30	61.2	448	2	Q9XVM6_CAEEL	Q9xvw6 caenorhabdi
905	30	61.2	340	2	Q26744_9TRYP	Q26744 trypanosoma	978	30	61.2	455	2	Q4K134_STRPN	Q4k134 streptococc
906	30	61.2	348	1	LIPA_CORGL	Q8nnj0 corynebacte	979	30	61.2	455	2	Q8KC87_CHLTE	Q8kc87 chlorobium
907	30	61.2	349	1	LIPA_CORDI	P61194 corynebacte	980	30	61.2	459	2	Q8NIB5_TALEM	Q8nib5 talaromyces

```
981 30 61.2 459 2 Q6ST59_9PICO Q6st59 human rhino
982 30 61.2 462 2 Q5QNU2_MOUSE Q5qnu2 mus musculus
983 30 61.2 463 1 VCL VICFA P08438 vicia faba
984 30 61.2 466 2 Q9V2D8_PYRAB Q9v2d8 pyrococcus
985 30 61.2 466 2 Q4QDW8_LEIMA Q4gdw8 leishmania
986 30 61.2 467 2 Q6LDG1_CAEBR Q6ldg1 caenorhabdi
987 30 61.2 469 2 Q6UNX5_NICBE Q6unx5 nicotiana b
988 30 61.2 470 2 Q9X1N3_THEMA Q9xin3 thermotoga
989 30 61.2 473 2 Q89JN2_BRAJA Q89jn2 bradyrhizob
990 30 61.2 481 2 Q9UWX2_SULSO Q9uwx2 sulfolobus
991 30 61.2 481 2 Q6BP54_DEBHA Q6bp54 debaryomyce
992 30 61.2 483 1 UXAB YERPE Q8zic5 yersinia pe
993 30 61.2 483 2 Q665N9_YERPS Q665n9 yersinia ps
994 30 61.2 489 2 Q9HMJ6_HALSA Q9hmj6 halobacteri
995 30 61.2 490 2 Q9W6G5_BRARE Q9w6g5 brachydanio
996 30 61.2 496 2 Q9HIK3_THEAC Q9hik3 thermoplasm
997 30 61.2 497 2 Q6L101_TRYCR Q6l101 trypanosoma
998 30 61.2 501 2 Q94450_SCHPO Q94450 schizosacch
999 30 61.2 505 2 Q8IA46_SPOFR Q8ia46 spodoptera
1000 30 61.2 505 2 Q8IU00_CAEEL Q8iu00 caenorhabdi
```

ALIGNMENTS

```
RESULT 1
Q4S6F3 TETNG
ID Q4S6F3_TETNG PRELIMINARY; PRT; 458 AA.
AC Q496F3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCAF14728, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0023316001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
```

```
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014728; CAG03779.1; -; Genomic_DNA.
FT NON TER 458 458
SQ SEQUENCE 458 AA; 51958 MW; 088CDE5D6EE5F977 CRC64;
```

Query Match 83.7%; Score 41; DB 2; Length 458;
Best Local Similarity 70.0%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ESVREQVMKY 10
Db 307 ESIREQIKKY 316
||:||||: ||
||:||||: ||

RESULT 2
Q9ZJN3_HELPUJ
ID Q9ZJN3_HELPUJ PRELIMINARY; PRT; 1164 AA.
AC Q9ZJN3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative.
GN OrderedLocusNames=JHP1272;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001550; AAD06844.1; -; Genomic_DNA.
DR PIR; G71827; G71827.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR PRINTS; PR00507; N12N6MTFRASE.
KW Complete proteome.
SQ SEQUENCE 1164 AA; 134378 MW; 5EC6286EA439873E CRC64;

Query Match 81.6%; Score 40; DB 2; Length 1164;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ESVREQVMKY 10
Db 740 ESLKEQVLKY 749
||:||||: ||
||:||||: ||

RESULT 3
Q4SFR7_TETNG
ID Q4SFR7_TETNG PRELIMINARY; PRT; 1064 AA.
AC Q4SFR7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 7 SCAF14601, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00018981001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
```


RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014601; CAG00515.1; -; Genomic_DNA.
FT NON_TER 1064 1064
SQ SEQUENCE 1064 AA; 116866 MW; AF3106816A29D570 CRC64;

Query Match 79.6%; Score 39; DB 2; Length 1064;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:||||:|
Db 986 ESIREHVIKY 995

RESULT 4
ERA_MYCPN STANDARD; PRT; 291 AA.
ID P75210;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE GTP-binding protein era homolog.
GN Name=era; Synonyms=spg; OrderedLocusNames=MPN568; ORFNames=MP274;
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;
RA Himmelreich R., Hilbert H., Plagens H., Pirkle E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
CC and is essential for cell growth (By similarity).
CC -!- SIMILARITY: Belongs to the era/trmE GTP-binding protein family.
CC Era subfamily.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE000027; AAB95922.1; -; Genomic_DNA.
DR PIR; S73600; S73600.
DR HAMAP; MF_00367; -; 1.
DR InterPro; IPR005662; Era GTP bd.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR004044; KH_TYPE 2.
DR InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR Pfam; PF07650; KH 2; 1.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OBG.
DR TIGRFAMs; TIGR00436; era; 1.
DR PROSITE; PS50823; KH_TYPE_2; 1.

KW Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
FT DOMAIN 201 279 KH type-2.
FT NP_BIND 10 17 GTP (Potential).
FT NP_BIND 58 62 GTP (Potential).
FT NP_BIND 122 125 GTP (Potential).
SQ SEQUENCE 291 AA; 33529 MW; 2E81B0406A9106E3 CRC64;

Query Match 77.6%; Score 38; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 79;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:||||:|
Db 191 EALREQIIKY 200

RESULT 5
Q52A99_MAGGR PRELIMINARY; PRT; 592 AA.
ID Q52A99_MAGGR PRELIMINARY;
AC Q52A99;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG02135.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasany U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
KW EMBL; AACU01000408; EAA54150.1; -; Genomic_DNA.
SQ SEQUENCE 592 AA; 68710 MW; 68F2DEBFBBCB9EC CRC64;

Query Match 77.6%; Score 38; DB 2; Length 592;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | : | | |
Db 540 EFVRERIMKY 549

RESULT 6
Q5CP31 CRYHO
ID Q5CP31 CRYHO PRELIMINARY; PRT; 593 AA.
AC Q5CP31;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE IFP53.
GN ORFNames=Chro.70176;
OS Cryptosporidium hominis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=237895;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=7U502;
RA Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G.,
RA Puiu D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H.,
RA Bankier A.T., Peterson D.L., Abrahamsen M.S., Kapur V., Tzipori S.,
RA Buck G.A.;
RT "The genome of Cryptosporidium hominis.";
RL Nature 431:1107-1112(2004).
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophyl-tRNA(Trp).
CC -!- SUBUNIT: Homodimer (By similarity).
DR EMBL; AAE01000003; EAL38376.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:Ligase activity; IEA.
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trpS; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
KW Protein biosynthesis.
SQ SEQUENCE 593 AA; 67816 MW; E2FECA7A13470D8E CRC64;

Query Match 77.6%; Score 38; DB 2; Length 593;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | : | | |

Db 483 ESIRNKIMKY 492

RESULT 7
Q5CYP8 CRYPV
ID Q5CYP8 CRYPV PRELIMINARY; PRT; 593 AA.
AC Q5CYP8;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Tryptophanyl-tRNA synthetase.
GN ORFNames=cgd7_1490;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RX PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamsen M.S., Templeton T.J., Enomoto S., Abrahante J.E., Zhu G.,
RA Lancto C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
RT parvum.";
RL Science 304:441-445(2004).
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophyl-tRNA(Trp).
CC -!- SUBUNIT: Homodimer (By similarity).
DR EMBL; AAE01000001; EAK90203.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:Ligase activity; IEA.
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1b.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trpS; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
KW Protein biosynthesis.
SQ SEQUENCE 593 AA; 67886 MW; ECE9587B2ECF6ECC CRC64;

Query Match 77.6%; Score 38; DB 2; Length 593;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | : | | |
Db 483 ESIRNKIMKY 492

RESULT 8
Q6JIZ0 MOUSE
ID Q6JIZ0 MOUSE PRELIMINARY; PRT; 2448 AA.
AC Q6JIZ0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyl-CoA carboxylase 2.
GN Name=Acacb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Heart;

RA Mao J., Wakil S.J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY451394; AAS13686.1; -, mRNA.
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR005482; Biotin carb C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; Cphp_synth_L_D2.
DR InterPro; IPR002114; HPr_Serp_S.
DR Pfam; PF02785; Biotin_lipoyl; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; Carboxyl_trans; 1.
DR Pfam; PF02786; CPsase_L_D2; 1.
DR PROSITE; PS50975; ATP_GRASP; 1.
DR PROSITE; PS50979; BC; 1.
DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
DR PROSITE; PS50989; COA_CT_CTER; 1.
DR PROSITE; PS50980; COA_CT_NTER; 1.
DR PROSITE; PS00866; CPsase_1; UNKNOWN 1.
DR PROSITE; PS00867; CPsase_2; UNKNOWN 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN 1.
SQ SEQUENCE 2448 AA; 275666 MW; 0B8649F5D2CF1C8A CRC64;

Query Match 77.6%; Score 38; DB 2; Length 2448;
Best Local Similarity 70.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | : | : |
Db 1599 ESVRDMVMRY 1608

RESULT 9
Q9LJZ1 ARATH PRELIMINARY; PRT; 394 AA.
ID Q9LJZ1 ARATH
AC Q9LJZ1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similarity to cytochrome P450.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221 (2000).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AP000383; BAB01867.1; -; Genomic DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.

DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00465; EP450IV.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 394 AA; 44658 MW; 5A3AAD918246BBA4 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : | | | : | |
Db 316 EEIREQALKY 325

RESULT 10
IF5_MOUSE
ID IF5_MOUSE STANDARD; PRT; 429 AA.
AC P59325;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN Name=Eif5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6, and FVB/N; TISSUE=Colon, and Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Catalyzes the hydrolysis of GTP bound to the 40S
CC ribosomal initiation complex (40S.mRNA.Met-tRNA[^{Phe}].eIF-2.GTP) with
CC the subsequent joining of a 60S ribosomal subunit resulting in the
CC release of eIF-2 and the guanine nucleotide. The subsequent
CC joining of a 60S ribosomal subunit results in the formation of a
CC functional 80S initiation complex (80S.mRNA.Met-tRNA[^{Phe}]).
CC -!- SIMILARITY: Belongs to the eIF-2-beta/eIF-5 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BC039275; AAH39275.1; -; mRNA.
DR EMBL; BC042622; AAH42622.1; -; mRNA.
DR Ensembl; ENSMUSG00000021282; Mus musculus.


```
DR MGI:95309; Eif5.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
KW GTP-binding; Initiation factor; Nucleotide-binding; Phosphorylation;
KW Protein biosynthesis.
FT NP_BIND 27 34 GTP (Potential).
FT COMPBIAS 194 200 Asp/Glu-rich (highly acidic).
FT COMPBIAS 382 400 Asp/Glu-rich (highly acidic).
FT COMPBIAS 421 427 Asp-rich (acidic).
FT MOD_RES 387 387 Phosphoserine (By similarity).
FT MOD_RES 388 388 Phosphoserine (By similarity).
SQ SEQUENCE 429 AA; 48968 MW; C77DD5DFBB5CIDEF CRC64;

Query Match 75.5%; Score 37; DB 1; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
|:|:|:|

RESULT 11
IF5_RAT STANDARD; PRT; 429 AA.
AC Q07205;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN Name=Eif5;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 131-150; 215-242 AND
RP 339-353.
RC TISSUE=Liver;
RX MEDLINE=93219424; PubMed=8464924;
RA Das K., Chevesich J., Maitra U.;
RT "Molecular cloning and expression of cDNA for mammalian translation
RT initiation factor 5.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:3058-3062(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PHOSPHORYLATION SITES SER-387 AND SER-388.
RX MEDLINE=21851300; PubMed=11861906; DOI=10.1093/nar/30.5.1154;
RA Majumdar R., Bandyopadhyay A., Deng H., Maitra U.;
RT "Phosphorylation of mammalian translation initiation factor 5 (eIF5)
RT in vitro and in vivo.";
RL Nucleic Acids Res. 30:1154-1162(2002).
CC -!- FUNCTION: Catalyzes the hydrolysis of GTP bound to the 40S
CC ribosomal initiation complex (40S.mRNA.Met-tRNA[F].eIF-2.GTP) with
CC the subsequent joining of a 60S ribosomal subunit resulting in the
CC release of eIF-2 and the guanine nucleotide. The subsequent
CC joining of a 60S ribosomal subunit results in the formation of a
CC functional 80S initiation complex (80S.mRNA.Met-tRNA[F]).
CC -!- SUBUNIT: Monomer; but some authors state that this protein may not
CC be a monomer.
CC -!- SIMILARITY: Belongs to the eIF-2-beta/eIF-5 family.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; L11651; AAA41112.1; -; mRNA.
DR EMBL; BC062398; AAH62398.1; -; mRNA.
DR PIR; A47305; A47305.
DR Ensembl; ENSRNOG00000010218; Rattus norvegicus.
DR RGD; 619861; Eif5.
DR GO; GO:0003743; F:translation initiation factor activity; IDA.
DR GO; GO:0006412; P:protein biosynthesis; TAS.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
KW Direct protein sequencing; GTP-binding; Initiation factor;
KW Nucleotide-binding; Phosphorylation; Protein biosynthesis.
FT NP_BIND 27 34 GTP (Potential).
FT COMPBIAS 194 200 Asp/Glu-rich (highly acidic).
FT COMPBIAS 382 400 Asp/Glu-rich (highly acidic).
FT COMPBIAS 421 427 Asp-rich (acidic).
FT MOD_RES 387 387 Phosphoserine.
FT MOD_RES 388 388 Phosphoserine.
SQ SEQUENCE 429 AA; 48954 MW; B1A62E30936908EE CRC64;

Query Match 75.5%; Score 37; DB 1; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
|:|:|:|

RESULT 12
Q5ZIE0 CHICK
ID Q5ZIE0_CHICK PRELIMINARY; PRT; 430 AA.
AC Q5ZIE0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_27j21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ720844; CAG32503.1; -; mRNA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006446; P:regulation of translational initiation; IEA.
DR InterPro; IPR003307; eIF5C.
DR InterPro; IPR002735; eIF5_eIF2B.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
KW Hypothetical protein.
SQ SEQUENCE 430 AA; 49078 MW; 6105AC7230CEA878 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 430;
```

Best Local Similarity 60.0%; Pred. No. 1.8e+02;									
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;									
QY	1 ESVREQVMKY 10								
	: : :								
Db	285 EKIREQIRKY 294								
RESULT 13									
IF5_HUMAN									
ID	IF5_HUMAN	STANDARD;	PRT;	431 AA.					
AC	P55010; Q9H5N2; Q9UG48;								
DT	01-OCT-1996 (Rel. 34, Created)								
DT	28-FEB-2003 (Rel. 41, Last sequence update)								
DT	13-SEP-2005 (Rel. 48, Last annotation update)								
DE	Eukaryotic translation initiation factor 5 (eIF-5).								
GN	Name=EIF5;								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;								
OC	Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	NUCLEOTIDE SEQUENCE.								
RX	MEDLINE=96279275; PubMed=8663286; DOI=10.1074/jbc.271.28.16934;								
RA	Si K., Das K., Maitra U.;								
RT	"Characterization of multiple mRNAs that encode mammalian translation								
RT	initiation factor 5 (eIF-5).";								
RL	J. Biol. Chem. 271:16934-16938(1996).								
RN	[2]								
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].								
RC	TISSUE=Brain;								
RX	MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;								
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,								
RA	Ansorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,								
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,								
RA	Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,								
RA	Wambutt R., Korn B., Klein M., Poustka A.;								
RT	"Towards a catalog of human genes and proteins: sequencing and								
RT	analysis of 500 novel complete protein coding human cDNAs.";								
RL	Genome Res. 11:422-435(2001).								
RN	[3]								
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].								
RX	PubMed=14702039; DOI=10.1038/ng1285;								
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,								
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,								
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,								
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,								
RA	Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,								
RA	Shiratori A., Sudoh H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,								
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,								
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,								
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,								
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,								
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,								
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,								
RA	Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,								
RA	Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,								
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,								
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,								
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,								
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,								
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,								
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,								
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,								
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,								
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,								
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,								
RA	Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,								
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,								
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,								
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;								
RT	"Complete sequencing and characterization of 21,243 full-length human								

RT	CDNAs.";						
RL	Nat. Genet. 36:40-45(2004).						
RN	[4]						
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].						
RC	TISSUE=Brain, and Uterus;						
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;						
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,						
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,						
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,						
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,						
RA	Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,						
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,						
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,						
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,						
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,						
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,						
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,						
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,						
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,						
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,						
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,						
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,						
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;						
RT	"Generation and initial analysis of more than 15,000 full-length human						
RT	and mouse cDNA sequences.";						
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).						
CC	-!- FUNCTION: Catalyzes the hydrolysis of GTP bound to the 40S						
CC	ribosomal initiation complex (40S.mRNA.Met-tRNA[F].eIF-2.GTP) with						
CC	the subsequent joining of a 60S ribosomal subunit resulting in the						
CC	release of eIF-2 and the guanine nucleotide. The subsequent						
CC	joining of a 60S ribosomal subunit results in the formation of a						
CC	functional 80S initiation complex (80S.mRNA.Met-tRNA[F]).						
CC	-!- INTERACTION:						
CC	O14602:EIF1AY; NbExp=1; IntAct=EBI-286450, EBI-286439;						
CC	-!- SIMILARITY: Belongs to the eIF-2-beta/eIF-5 family.						
CC	-----						
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use as long as its content is in no way modified and this statement is not						
CC	removed.						
CC	-----						
DR	EMBL; U49436; AAC50572.1; -; mRNA.						
DR	EMBL; AL080102; CAB45711.1; -; mRNA.						
DR	EMBL; AK026933; BAB15593.1; -; mRNA.						
DR	EMBL; BC007728; AAH07728.1; -; mRNA.						
DR	EMBL; BC032866; AAH32866.1; -; mRNA.						
DR	PIR; T12450; T12450.						
DR	IntAct; P55010; -.						
DR	Ensembl; ENSG00000100664; Homo sapiens.						
DR	HGNC; HGNC:3299; EIF5.						
DR	H-InvDB; HIX001199; -.						
DR	Reactome; P55010; -.						
DR	MIM; 601710; -.						
DR	GO; GO:0005829; C:cytosol; NAS.						
DR	GO; GO:0003924; F:GTPase activity; TAS.						
DR	GO; GO:0003743; F:translation initiation factor activity; NAS.						
DR	GO; GO:0006446; P:regulation of translational initiation; TAS.						
DR	InterPro; IPR002735; eIF5 eIF2B.						
DR	InterPro; IPR003307; eIF5C.						
DR	Pfam; PF01873; eIF-5 eIF-2B; 1.						
DR	Pfam; PF02020; W2; 1.						
DR	ProDom; PD004078; eIF5 eIF2B; 1.						
DR	SMART; SM00653; eIF2B_5; 1.						
DR	SMART; SM00515; eIF5C; 1.						
DR	GTP-binding; Initiation factor; Nucleotide-binding; Phosphorylation;						
KW	Protein biosynthesis.						
KW	NP BIND 27 34 GTP (Potential).						
FT	COMPBIAS 196 202 Asp/Glu-rich (highly acidic).						
FT	COMPBIAS 384 402 Asp/Glu-rich (highly acidic).						
FT	COMPBIAS 423 429 Asp-rich (acidic).						
FT	MOD_RES 389 389 Phosphoserine (By similarity).						
FT	MOD_RES 390 390 Phosphoserine (By similarity).						

FT CONFLICT 60 60 E -> G (in Ref. 2).
FT CONFLICT 203 203 W -> S (in Ref. 1).
FT CONFLICT 295 295 K -> E (in Ref. 2).
FT CONFLICT 311 311 Q -> K (in Ref. 4; AAH32866).
SQ SEQUENCE 431 AA; 49223 MW; C6CCC3A255F9B9BC CRC64;

Query Match 75.5%; Score 37; DB 1; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 14
Q53XB3 HUMAN PRELIMINARY; PRT; 431 AA.
AC Q53XB3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein DKFZp686K0727.
GN Name=DKFZp686K0727;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Human testis;
RA Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX537367; CAD97610.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 49222 MW; C6CCC3A255F9B9BC CRC64;

Query Match 75.5%; Score 37; DB 2; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 15
Q5R4L0 PONPY PRELIMINARY; PRT; 431 AA.
AC Q5R4L0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459F153.
GN Name=DKFZp459F153;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German CDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR861236; CAH93306.1; -; mRNA.
DR GO; GO:0003743; P:translation initiation factor activity; IEA.
DR GO; GO:0006446; P:regulation of translational initiation; IEA.
DR InterPro; IPR003307; eIF5C.

DR InterPro; IPR002735; eIF5 eIF2B.
DR Pfam; PF01873; eIF-5 eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5 eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 49211 MW; 9789C6C01D7638CC CRC64;

Query Match 75.5%; Score 37; DB 2; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 16
Q7QPI9 GIALA PRELIMINARY; PRT; 892 AA.
AC Q7QPI9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 173 13416 16094.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000168; EAA36927.1; -; Genomic_DNA.
DR InterPro; IPR002553; Adaptin_N.
DR Pfam; PF01602; Adaptin_N; 1.
SQ SEQUENCE 892 AA; 99003 MW; 9E78811CD1B98189 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 892;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 593 ESARKQVMEY 602

RESULT 17
Q8I5C3 PLAF7 PRELIMINARY; PRT; 1132 AA.
AC Q8I5C3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PFL1430C;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,


```
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Tallia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confariolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
CC -1- SIMILARITY: Belongs to the ribosomal protein S15P family.
DR EMBL; CR380953; CAG59506.1; -; Genomic_DNA.
DR SMR; Q6FT65; 66-130.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR00589; Ribosomal S15.
DR Pfam; PF00312; Ribosomal S15; 1.
DR PROSITE; PS00362; RIBOSOMAL_S15; 1.
KW Complete proteome; Ribonucleoprotein; Ribosomal protein.
SQ SEQUENCE 151 AA; 17061 MW; 5594AFD07342ADF7 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 151;
Best Local Similarity 70.0%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
   ||| |||:|
Db 31 ESVIEQIIKY 40

RESULT 20
YBIA_ECOLI STANDARD; PRT; 160 AA.
AC P30176;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical protein ybIA.
GN Name=ybIA; OrderedLocusNames=b0798;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12 / W3110;
RX MEDLINE=94312114; PubMed=8037924;
RA Ohmori H.;
RT "Structural analysis of the rhlE gene of Escherichia coli.";
RL Jpn. J. Genet. 69:1-12(1994).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
```

```
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sanpei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L02123; AAA53654.1; -; Genomic DNA.
DR EMBL; U00096; AAC73885.1; -; Genomic_DNA.
DR EMBL; D90716; BAA35458.1; -; Genomic_DNA.
DR EMBL; D90717; BAA35464.1; -; Genomic_DNA.
DR PIR; F64816; F64816.
DR ECHOBASE; EB1539; -.
DR EcoGene; EG11579; ybIA.
DR InterPro; IPR012816; CHP02464_YbIA.
DR TIGRFAMs; TIGR02464; ribofla_fusion; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 160 AA; 18669 MW; 28D6DA5C6748653C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 160;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
   |||:|
Db 90 ESVKEQVMR 98

RESULT 21
Q83LU8_SHIFL PRELIMINARY; PRT; 160 AA.
ID Q83LU8_SHIFL PRELIMINARY;
AC Q83LU8; Q7C2G6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein ybIA.
GN Name=ybIA; OrderedLocusNames=S0789, SF0749;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
```

```
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AB005674; AAN42383.1; -; Genomic_DNA.
DR EMBL; AB016980; AAP16260.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 160 AA; 18629 MW; AB0015667247722A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 160;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
Db 90 ESVKEQVMR 98

RESULT 22
Q50HZ9 BOVIN
ID Q50HZ9_BOVIN PRELIMINARY; PRT; 187 AA.
AC Q50HZ9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Acetyl-CoA carboxylase, type beta (Fragment).
GN Name=acacib;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart muscle;
RA Lovstedt K., Hallen S., Osla F., Clapham J., Corneliussen B.;
RT "Isolation and characterization of cDNA encoding bovine heart acetyl-
RT CoA carboxylase type beta.";
RL Submitted (APR-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ966323; CAI84636.1; -; mRNA.
FT NON_TER 1
FT NON_TER 187
SQ SEQUENCE 187 AA; 21758 MW; F30C011A2705CBB5 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 187;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 52 ESVRSMVMRY 61

RESULT 23
Q4P3L0 USTMA
ID Q4P3L0_USTMA PRELIMINARY; PRT; 193 AA.
AC Q4P3L0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=UM05303.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
```

```
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseis M., Karlsson E.,
RA Kellis C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP0100191; EAK86552.1; -; Genomic_DNA.
SQ SEQUENCE 193 AA; 21243 MW; 657A60F0285CF8A6 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 193;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 85 QSIREMLKY 94

RESULT 24
Q62DC6 BURMA
ID Q62DC6_BURMA PRELIMINARY; PRT; 242 AA.
AC Q62DC6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Phosphoglycolate phosphatase (EC 3.1.3.18).
GN Name=gph-2; OrderedLocusNames=BMAA0534;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
```


RA Madupu R., Mohammad Y., Nelson W.C., Radune D., Romero C.M., Yu Y.,
RA Sarria S., Selengut J., Shamblyn C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000011; AAU46883.1; -; Genomic_DNA.
DR TIGR; BMAA0534; -;
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR006402; HAD_SF-IA-v3.
DR InterPro; IPR006439; HAD_SF-IA-v1.
DR InterPro; IPR005833; Hlg_nase/ep_hydro.
DR InterPro; IPR006346; PGP_bact.
DR Pfam; PF00702; Hydrolase_1.
DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-v3; 1.
DR TIGRFAMs; TIGR01449; PGP_bact; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 242 AA; 25312 MW; 652EFBA29619ABD5 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 242;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 45 ETTREVMRY 54

RESULT 25
Q63QH4_BURPS
ID Q63QH4_BURPS PRELIMINARY; PRT; 245 AA.
AC Q63QH4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative hydrolase.
GN OrderedLocusNames=BPSL3049;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crossett B., Davis P., DeShazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchaveest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH37060.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR006402; HAD_SF-IA-v3.
DR InterPro; IPR006439; HAD_SF_A_v1.
DR InterPro; IPR005833; Hlg_nase/ep_hydro.
DR InterPro; IPR006346; PGP_bact.
DR Pfam; PF00702; Hydrolase; 1.

DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-v3; 1.
DR TIGRFAMs; TIGR01449; PGP_bact; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 245 AA; 25602 MW; 3FDEE0CE60E884E33 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 245;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 48 ETTREVMRY 57

RESULT 26
Q6A7W0_PROAC
ID Q6A7W0_PROAC PRELIMINARY; PRT; 418 AA.
AC Q6A7W0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tyrosyl-tRNA synthetase (EC 6.1.1.1).
GN OrderedLocusNames=PPA1404;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacterineae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hoster F., Liesegang H., Wiewer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
DR EMBL; AB017283; AAT83155.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002942; S4.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; Tyr tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTHTYR.
DR TIGRFAMs; TIGR00234; tyrS; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS50889; S4; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
SQ SEQUENCE 418 AA; 46257 MW; 391E0973BAD5E707 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 418;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|:
Db 99 ESIKDQVSKY 108

RESULT 27
Q9FH67_ARATH
ID Q9FH67_ARATH PRELIMINARY; PRT; 499 AA.
AC Q9FH67;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome P450.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AB022210; BAB09329.1; -; Genomic_DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 499 AA; 56882 MW; B2E7BA86E4348397 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 499;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| :||| :||
Db 420 EKIREQAVKY 429

RESULT 28
Q9LIG9 ARATH
ID Q9LIG9_ARATH PRELIMINARY; PRT; 523 AA.
AC Q9LIG9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome P450-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AP001304; BAB01904.1; -; Genomic_DNA.
DR HSSP; P14779; IJME.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.

DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 523 AA; 59570 MW; F1D4BEF899D60E7A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 523;
Best Local Similarity 60.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
: |||: :||
Db 436 DEVREEVLKY 445

RESULT 29
Q529X0 MAGGR
ID Q529X0_MAGGR PRELIMINARY; PRT; 671 AA.
AC Q529X0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG02264.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-chomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Testaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000439; EAA54279.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 671 AA; 77945 MW; FD24B32542D6AA78 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 671;
Best Local Similarity 70.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | |
Db 602 EFVRERMMKY 611

RESULT 30
Q51V60 MAGGR PRELIMINARY; PRT; 829 AA.
AC Q51V60;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG04275.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysseilis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoswo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,

RA Towe S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001087; EAA50516.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 829 AA; 96375 MW; 97C6F404C36235F8 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 829;
Best Local Similarity 70.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | |
Db 740 EFVRERMMKY 749

RESULT 31
Q5CWC1 CRYPV PRELIMINARY; PRT; 970 AA.
AC Q5CWC1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Ydr449cp/Utp6p; small (Ribosomal) subunit (SSU) processosome (Contains
DE U3 snRNA). HAT repeats.
GN ORFNames=cgd8_900;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RX PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahante J.E., Zhu G.,
RA Lanceto C.A., Deng M., Liu C., Widmer G., Tripori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
RT parvum."
RL Science 304:441-445(2004).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAEE01000003; EAK89317.1; -; Genomic_DNA.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR003164; AP2_A_adaptin_C.
DR InterPro; IPR008456; Collagen_bind.
DR InterPro; IPR011992; EF-Hand_type.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001093; IMPDh/GMPrtase.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR000403; PI3/4_kinase_cat.
DR InterPro; IPR008925; tRNA-synt_bind.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00478; IMPDH; 1.

JARD2_MOUSE
ID JARD2_MOUSE STANDARD; PRT; 1234 AA.
AC Q62315; Q99LD1;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Jumonji protein (Jumonji/ARID domain-containing protein 2).
GN NamesJarid2; Synonyms=Jmj;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=129/Ola;
RX MEDLINE=95278734; PubMed=7758946;
RA Takeuchi T., Yamazaki Y., Katoh-Fukui Y., Tsuchiya R., Kondo S.,
RA Motoyama J., Higashinakagawa T.;
RT "Gene trap capture of a novel mouse gene, jumonji, required for neural
RT tube formation.";
RL Genes Dev. 9:1211-1222(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6, and Czech II; TISSUE=Brain, and Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=20374468; PubMed=10913339; DOI=10.1006/bbrc.2000.3138;
RA Toyoda M., Kojima M., Takeuchi T.;
RT "Jumonji is a nuclear protein that participates in the negative
RT regulation of cell growth.";
RL Biochem. Biophys. Res. Commun. 274:332-336(2000).
RN [5]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=20269961; PubMed=10807864;
RA Lee Y., Song A.J., Baker R., Micales B., Conway S.J., Lyons G.E.;
RT "Jumonji, a nuclear protein that is necessary for normal heart
RT development.";
RL Circ. Res. 86:932-938(2000).
RN [6]
RP FUNCTION, NUCLEAR LOCALIZATION SIGNAL, AND MUTAGENESIS OF
RX 106-ARG-LYS-107.
RX PubMed=12890668; DOI=10.1074/jbc.M307386200;
RA Kim T.-G., Kraus J.C., Chen J., Lee Y.;
RT "JUMONJI, a critical factor for cardiac development, functions as a
RT transcriptional repressor.";
RL J. Biol. Chem. 278:42247-42255(2003).
RN [7]
RP FUNCTION, AND INTERACTIONS WITH GATA4 AND NKX2-5.
RX PubMed=15542826; DOI=10.1128/MCB.24.23.10151-10160.2004;
RA Kim T.-G., Chen J., Sadoshima J., Lee Y.;
RT "Jumonji represses atrial natriuretic factor gene expression by
RT inhibiting transcriptional activities of cardiac transcription
RT factors.";
RL Mol. Cell. Biol. 24:10151-10160(2004).
CC -!- FUNCTION: Required for neural tube formation. Essential for normal
CC heart development and function. Acts as a transcriptional
CC repressor of ANF by binding to both GATA4 and NKX2-5 and
CC repressing their transcriptional activator activities.
CC Participates in the negative regulation of cell proliferation
CC signaling.
CC -!- SUBUNIT: Interacts with the N-terminal region of GATA4 and the C-
CC terminal region of the NKX2-5 homeobox and represses their ability
CC to activate transcription of ANF.
CC -!- INTERACTION:
CC Q08369:Gata4; NbExp=2; IntAct=EBI-493592, EBI-297008;
CC P42582:Nkx2-5; NbExp=2; IntAct=EBI-493592, EBI-297021;
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Widely expressed in embryos. In adults,
CC expressed at high levels in heart, skeletal muscle, brain and
CC thymus.
CC -!- SIMILARITY: Contains 1 ARID domain.
CC -!- SIMILARITY: Contains 1 JmjC domain.
CC -!- SIMILARITY: Contains 1 JmjN domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; D31967; BAA06736.1; -; mRNA.
CC EMBL; AK045214; BAC32264.1; -; mRNA.
CC EMBL; BC003374; AAH03374.1; -; mRNA.
CC EMBL; BC052444; AAH52444.1; -; mRNA.
CC EMBL; BC060695; AAH60695.1; -; mRNA.
CC PIR; T30254; T30254.
CC IntAct; Q62315; -.
CC Ensembl; ENSMUSG00000038518; Mus musculus.
CC MGI; MGI:104813; Jarid2.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003677; F:DNA binding; IDA.

DR GO; GO:0016564; F:transcriptional repressor activity; IDA.
DR GO; GO:0045892; P:negative regulation of transcription, DNA-d. . .; IDA.
DR InterPro; IPR001606; ARID.
DR InterPro; IPR003347; TF_JmjC.
DR InterPro; IPR003349; TF_JmjN.
DR InterPro; IPR004198; ZnF_C5HC2.
DR Pfam; PF01388; ARID; 1.
DR Pfam; PF02373; JmjC; 1.
DR Pfam; PF02375; JmjN; 1.
DR Pfam; PF02928; zf-C5HC2; 1.
DR PROSITE; PS51011; ARID; 1.
KW Developmental protein; Nuclear protein; Repressor; Transcription;
KW Transcription regulation.
FT DOMAIN 619 711 ARID.
FT MOTIF 104 110 Nuclear localization signal.
FT MUTAGEN 106 107 RK->AA: Leads to cytoplasmic
FT relocalization.
FT CONFLICT 1096 1096 L -> Q (in Ref. 3; AAH03374).
SQ SEQUENCE 1234 AA; 137445 MW; B56E172C5E5745B5 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 1234;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 573 ESVRAQVEKY 582
|||||

RESULT 36
Q86LZ8 SCHMA PRELIMINARY; PRT; 1499 AA.
AC Q86LZ8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Insulin receptor protein kinase RTK-2.
GN Name=RTK-2;
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22442601; PubMed=12554084; DOI=10.1016/S0166-6851(02)00249-9;
RA Vicogne J., Pin J.P., Lardans V., Capron M., Noel C., Dissois C.;
RT "An unusual receptor tyrosine kinase of Schistosoma mansoni contains a
RT Venus Flytrap module.";
RL Mol. Biochem. Parasitol. 126:51-62(2003).
DR EMBL; AF314754; AAN39120.1; -; mRNA.
DR HSSP; P08581; 1RIW.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR006212; Furin_repeat.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase; Receptor.

SQ SEQUENCE 1499 AA; 170354 MW; A603560611E26582 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 1499;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVREQVMKY 10
Db 176 SVREQIVRY 184
|||||

RESULT 37
Q8IDM0 PLAF7 PRELIMINARY; PRT; 1847 AA.
AC Q8IDM0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein MAL13P1.239.
GN Name=MAL13P1.239;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52601.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1847 AA; 221588 MW; DA999B6FE8DFF25A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 1847;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 854 ENTREQIMNY 863
|: |||:|

RESULT 38
COAC_CHICK STANDARD; PRT; 2324 AA.
ID COAC_CHICK
AC P11029;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Name=ACAC;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=88139305; PubMed=2893793;
RA Takai T., Yokoyama C., Wada K., Tanabe T.;
RT "Primary structure of chicken liver acetyl-CoA carboxylase deduced
RT from cDNA sequence.";
RL J. Biol. Chem. 263:2651-2657(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 493-820.
RC TISSUE=Liver;
RX MEDLINE=87106011; PubMed=2879745; DOI=10.1016/0014-5793(87)81564-8;
RA Takai T., Wada K., Tanabe T.;
RT "Primary structure of the biotin-binding site of chicken liver acetyl-
RT CoA carboxylase.";

RL FEBS Lett. 212:98-102(1987).
CC -|- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -|- COFACTOR: Biotin.
CC -|- ENZYME REGULATION: By phosphorylation.
CC -|- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Contains 1 ATP-grasp domain.
CC -|- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -|- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -|- SIMILARITY: Contains 1 carboxyltransferase domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC DR EMBL; J03541; AAA48701.1; -; mRNA.
CC DR EMBL; X05019; CAA28675.1; -; mRNA.
CC DR PIR; A29924; A29924.
CC DR HSSP; Q00955; I0D4.
CC DR Ensembl; ENSGALG0000005439; Gallus gallus.
CC DR InterPro; IPR011761; ATP_GRASP.
CC DR InterPro; IPR011764; BC.
CC DR InterPro; IPR001882; Biotin_BS.
CC DR InterPro; IPR005482; Biotin_carb_C.
CC DR InterPro; IPR000089; Biotin_lipoyl.
CC DR InterPro; IPR000022; Carboxyl_trans.
CC DR InterPro; IPR011763; COA_CT_C.
CC DR InterPro; IPR011762; COA_CT_N.
CC DR InterPro; IPR005481; CPase_L_N.
CC DR InterPro; IPR005479; CPase_D2_ATP_bd.
CC DR Pfam; PF02785; Biotin_carb_C; 1.
CC DR Pfam; PF03164; Biotin_lipoyl; 1.
CC DR Pfam; PF01039; Carboxyl_trans; 1.
CC DR Pfam; PF00289; CPase_L_chain; 1.
CC DR Pfam; PF02786; CPase_L_D2; 1.
CC DR PROSITE; PS50975; ATP_GRASP; 1.
CC DR PROSITE; PS50979; BC; 1.
CC DR PROSITE; PS00188; BIOTIN; 1.
CC DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
CC DR PROSITE; PS50989; COA_CT_CTER; 1.
CC DR PROSITE; PS50980; COA_CT_NTER; 1.
CC DR PROSITE; PS00866; CPASE_1; 1.
CC DR PROSITE; PS00867; CPASE_2; 1.
KW ATP-binding; Biotin; Direct protein sequencing;
KW Fatty acid biosynthesis; Ligase; Lipid synthesis;
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1675 2171 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).
FT ACT_SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent).
FT BINDING 1800 1800 Coenzyme A (By similarity).
FT BINDING 2104 2104 Coenzyme A (By similarity).
FT BINDING 2106 2106 Coenzyme A (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1193 1193 Phosphoserine (By similarity).
SQ SEQUENCE 2324 AA; 262720 MW; 3F1C541F01BBBF6 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2324;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
Db 1467 ESVRSMVMRY 1476
RESULT 39
COAL_RAT STANDARD; PRT; 2345 AA.
ID_COAL_RAT STANDARD; PRT; 2345 AA.
AC P11497; P97902;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Name=Acaca; Synonyms=Acac;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=88320328; PubMed=2901088;
RA Lopez-Casillas F., Bai D.-H., Luo X., Kong I.-S., Hermodson M.A.,
RA Kim K.-H.;
RT "Structure of the coding sequence and primary amino acid sequence of
RT acetyl-coenzyme A carboxylase.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5784-5788(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=89264558; PubMed=2566999;
RA Luo X.N., Park K., Lopez-Casillas F., Kim K.-H.;
RT "Structural features of the acetyl-CoA carboxylase gene: mechanisms
RT for the generation of mRNAs with 5' end heterogeneity.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4042-4046(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-33.
RX MEDLINE=89214151; PubMed=2565337;
RA Lopez-Casillas F., Kim K.-H.;
RT "Heterogeneity at the 5' end of rat acetyl-coenzyme A carboxylase
RT mRNA. Lipogenic conditions enhance synthesis of a unique mRNA in
RT liver.";
RL J. Biol. Chem. 264:7176-7184(1989).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1167-1200 (ISOFORMS 1 AND 2), AND
RP PHOSPHORYLATION SITE SER-1200.
RX MEDLINE=90337981; PubMed=1974251;
RA Kong I.-S., Lopez-Casillas F., Kim K.-H.;
RT "Acetyl-CoA carboxylase mRNA species with or without inhibitory coding
RT sequence for Ser-1200 phosphorylation.";
RL J. Biol. Chem. 265:13695-13701(1990).
RN [5]
RP BIOTIN-BINDING SITE.
RX MEDLINE=89289706; PubMed=2567668;
RA Bai D.-H., Moon T.-W., Lopez-Casillas F., Andrews P.C., Kim K.-H.;
RT "Analysis of the biotin-binding site on acetyl-CoA carboxylase from
RT rat.";
RL Eur. J. Biochem. 182:239-245(1989).
RN [6]
RP PROTEIN SEQUENCE OF 76-85 AND 1198-1201, AND PHOSPHORYLATION.
RX MEDLINE=88296498; PubMed=2900138;
RA Munday M.R., Campbell D.G., Carling D., Hardie D.G.;
RT "Identification by amino acid sequencing of three major regulatory
RT phosphorylation sites on rat acetyl-CoA carboxylase.";
RL Eur. J. Biochem. 175:331-338(1988).
RN [7]
RP PARTIAL PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=94237850; PubMed=7910165;
RA Winz R., Hess D., Aebersold R., Brownsey R.W.;
RT "Unique structural features and differential phosphorylation of the

```
RT 280-kDa component (isozyme) of rat liver acetyl-CoA carboxylase." ;
RL J. Biol. Chem. 269:14438-14445(1994).
CC -!- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- COFACTOR: Biotin.
CC -!- ENZYME REGULATION: By phosphorylation.
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P11497-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P11497-2; Sequence=VSP_011753;
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Contains 1 ATP-grasp domain.
CC -!- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -!- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -!- SIMILARITY: Contains 1 carboxyltransferase domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; J03808; AAA40653.1; -; mRNA.
CC EMBL; M26731; AAA40652.1; -; Genomic_DNA.
CC EMBL; M26195; AAA40654.1; -; mRNA.
CC EMBL; M26196; AAA40655.1; -; mRNA.
CC EMBL; M26197; AAA40656.1; -; mRNA.
CC EMBL; M55315; -; NOT_ANNOTATED_CDS; mRNA.
CC PIR; A35578; A35578.
CC HSSP; Q00955; LOD4.
CC Ensembl; ENSRNOG0000034013; Rattus norvegicus.
CC RGD; 621248; Acac.
CC InterPro; IPR011761; ATP_GRASP.
CC InterPro; IPR011764; BC.
CC InterPro; IPR001882; Biotin_BS.
CC InterPro; IPR005482; Biotin_carb_C.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR000022; Carboxyl_trans.
CC InterPro; IPR011763; COA_CT_C.
CC InterPro; IPR005481; CPase_L_N.
CC InterPro; IPR005479; CPase_D2_ATP_bd.
CC Pfam; PF02785; Biotin_carb_C; 1.
CC Pfam; PF00364; Biotin_lipoyl; 1.
CC Pfam; PF01039; Carboxyl_trans; 1.
CC Pfam; PF02786; CPase_L_D2; 1.
CC PROSITE; PS50975; ATP_GRASP; 1.
CC PROSITE; PS50979; BC; 1.
CC PROSITE; PS00188; BIOTIN; 1.
CC PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
CC PROSITE; PS50989; COA_CT_CTER; 1.
CC PROSITE; PS50980; COA_CT_NTER; 1.
CC PROSITE; PS00866; CPASE_1; 1.
CC PROSITE; PS00867; CPASE_2; 1.
KW Alternative splicing; ATP-binding; Biotin; Direct protein sequencing;
KW Fatty acid biosynthesis; Ligase; Lipid synthesis;
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 116 617 Biotin carboxylation.
FT DOMAIN 274 465 ATP-grasp.
FT DOMAIN 751 817 Biotinyl/lipoyl.
FT DOMAIN 1697 2193 Carboxyltransferase.
```

```
FT NP BIND 314 319 ATP (Potential).
FT ACT_SITE 440 440 By similarity.
FT BINDING 785 785 Biotin (covalent).
FT BINDING 1822 1822 Coenzyme A (By similarity).
FT BINDING 2126 2126 Coenzyme A (By similarity).
FT BINDING 2128 2128 Coenzyme A (By similarity).
FT MOD_RES 77 77 Phosphoserine.
FT MOD_RES 79 79 Phosphoserine.
FT MOD_RES 1200 1200 Phosphoserine.
FT VARSP_LIC 1189 1196 Missing (in isoform 2).
FT SQ SEQUENCE 2345 AA; 265194 MW; 78E9CF9ADE1E8771 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2345;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 1489' ESVRSMVMRY 1498

RESULT 40
Q5SWU9 MOUSE
ID Q5SWU9_MOUSE PRELIMINARY; PRT; 2345 AA.
AC Q5SWU9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Acetyl-Coenzyme A carboxylase.
GN Name=Acac; ORFNames=RP23-123010.1-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Matthews N.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Tracey A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596447; CAI25271.1; -; Genomic_DNA.
DR EMBL; AL596252; CAI24019.1; -; Genomic_DNA.
DR EMBL; AL596252; CAI25271.1; JOINED; Genomic_DNA.
DR EMBL; AL596447; CAI24019.1; JOINED; Genomic_DNA.
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005479; Cphs_synth_L_D2.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS50975; ATP_GRASP; 1.
DR PROSITE; PS50979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
```



```
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1698 2194 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).
FT ACT_SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent) (By similarity).
FT BINDING 1823 1823 Coenzyme A (By similarity).
FT BINDING 2127 2127 Coenzyme A (By similarity).
FT BINDING 2129 2129 Coenzyme A (By similarity).
FT MOD_RES 29 29 Phosphoserine (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1201 1201 Phosphoserine (By similarity).
SQ SEQUENCE 2346 AA; 265303 MW; 32886C5D03EEAE0E CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 1490 ESVRSMVMRY 1499

RESULT 43
COAL_HUMAN STANDARD; PRT; 2346 AA.
AC Q13085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Name=ACACA; Synonyms=ACAC, ACC1, ACCA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95249602; PubMed=7732023;
RA Abu-Elheiga L., Jayakumar A., Baldini A., Chirala S.S., Wakil S.J.;
RT "Human acetyl-CoA carboxylase: Characterization, molecular cloning,
RT and evidence for two isoforms.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:4011-4015(1995).
RN [2]
RX PHOSPHORYLATION SITE SER-29.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
RA Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC -!- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- COFACTOR: Biotin.
CC -!- ENZYME REGULATION: By phosphorylation (By similarity).
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Expressed in brain, placental, skeletal
CC muscle, renal, pancreatic and adipose tissues; expressed at low
CC level in pulmonary tissue; not detected in the liver.
CC -!- SIMILARITY: Contains 1 ATP-grasp domain.
```

```
CC -!- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -!- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -!- SIMILARITY: Contains 1 carboxyltransferase domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U19822; AAC50139.1; -; mRNA.
DR PIR; I38928; I38928.
DR HSSP; Q00955; IOD4.
DR Ensembl; ENSG00000132142; Homo sapiens.
DR HGNC; HGNC:84; ACACA.
DR MIM; 200350; -.
DR GO; GO:0003989; F:acetyl-CoA carboxylase activity; TAS.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; CPase_D2_ATP_bd.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS50975; ATP_GRASP; 1.
DR PROSITE; PS50979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
DR PROSITE; PS50989; COA_CT_CTER; 1.
DR PROSITE; PS50980; COA_CT_NTER; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW ATP-binding; Biotin; Fatty acid biosynthesis; Ligase; Lipid synthesis;
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1698 2194 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).
FT ACT_SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent) (By similarity).
FT BINDING 1823 1823 Coenzyme A (By similarity).
FT BINDING 2127 2127 Coenzyme A (By similarity).
FT BINDING 2129 2129 Coenzyme A (By similarity).
FT MOD_RES 29 29 Phosphoserine.
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1201 1201 Phosphoserine (By similarity).
SQ SEQUENCE 2346 AA; 265040 MW; 9519190D40190D14 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
Db 1490 ESVRSMVMRY 1499

RESULT 44
COAL_SHEEP STANDARD; PRT; 2346 AA.
ID COAL_SHEEP
AC Q28559;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
```

DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Names=ACACA; Synonyms=ACAC;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Finn-Dorset; TISSUE=Adipose tissue;
RX MEDLINE=95197015; PubMed=7890176; DOI=10.1016/0378-1119(94)00871-O;
RA Barber M.C., Travers M.T.;
RT "Cloning and characterisation of multiple acetyl-CoA carboxylase
RL transcripts in ovine adipose tissue.";
RL Gene 154:271-275(1995).
CC -!- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- COPACTOR: Biotin.
CC -!- ENZYME REGULATION: By phosphorylation (By similarity).
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 1 ATP-grasp domain.
CC -!- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -!- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -!- SIMILARITY: Contains 1 carboxyltransferase domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; X80045; CAA56352.1; -; mRNA.
DR HSSP; Q00955; 1OD4.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; CPase_D2_ATP_bd.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS50975; ATP_GRASP; 1.
DR PROSITE; PS50979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
DR PROSITE; PS50989; COA_CT_CTER; 1.
DR PROSITE; PS50980; COA_CT_NTER; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW ATP-binding; Biotin; Fatty acid biosynthesis; Ligase; Lipid synthesis;
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1698 2194 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).

FT ACT SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent) (By similarity).
FT BINDING 1823 1823 Coenzyme A (By similarity).
FT BINDING 2127 2127 Coenzyme A (By similarity).
FT BINDING 2129 2129 Coenzyme A (By similarity).
FT MOD_RES 29 29 Phosphoserine (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1201 1201 Phosphoserine (By similarity).
SQ SEQUENCE 2346 AA; 265251 MW; BCA010ADF6CD24EF CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| | | | | | | |
Db 1490 ESVRSMVMRY 1499

RESULT 45
Q7Z561_HUMAN PRELIMINARY; PRT; 2346 AA.
ID Q7Z561; AC Q7Z561;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Acetyl-CoA carboxylase 1.
GN Name=ACCL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Adipocyte;
RX MEDLINE=22709197; PubMed=12810950; DOI=10.1073/pnas.1332670100;
RA Mao J., Chirala S.S., Wakil S.J.;
RT "Human acetyl-CoA carboxylase 1 gene: presence of three promoters and
RL heterogeneity at the 5'-untranslated mRNA region.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7515-7520(2003).
DR EMBL; AY315627; AAP94122.1; -; mRNA.
DR HSSP; P24182; 1DV1.
DR Ensembl; ENSG00000132142; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; CPase_D2_ATP_bd.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS50975; ATP_GRASP; 1.
DR PROSITE; PS50979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS50968; BIOTINYL_LIPOYL; 1.
DR PROSITE; PS50989; COA_CT_CTER; 1.
DR PROSITE; PS50980; COA_CT_NTER; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
SQ SEQUENCE 2346 AA; 265615 MW; AC156BCBD15F77AB CRC64;

RL Submitted (NOV-2004) to the EMBL/GenBank/DBDJ databases.
DR EMBL; CR858529; CAH90756.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 74 AA; 8507 MW; F3C52A7E398AD18B CRC64;

Query Match 71.4%; Score 35; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ESVREQVMKY 10
|||:||||
Db 48 ESVKEQEMKW 57

RESULT 49

Q8R034_MOUSE
ID Q8R034_MOUSE PRELIMINARY; PRT; 74 AA.
AC Q8R034;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Anapc13 protein (Mus musculus 10 day old male pancreas cDNA, RIKEN
full-length enriched library, clone:1810004D07 product:HYPOTHETICAL
8.5 kDa PROTEIN homolog).
GN Name=Anapc13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Vexardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).

```

[8] NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Pancreas;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC028526; AAH28526.1; -; mRNA.
DR EMBL; AK007332; BAC25171.1; -; mRNA.
DR Ensemble; ENSMUSG00000035048; Mus musculus.
DR MGI; MGI:1916260; Anapc13.
KW Hypothetical protein.
SQ SEQUENCE 74 AA; 8346 MW; E84F16A13821CB3B CRC64;
```

```

Query Match 71.4%; Score 35; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
   |||:||||:
DB 48 ESVKEQEMKW 57
```

```

RESULT 50
Q4SKC6_TETNG
ID Q4SKC6_TETNG PRELIMINARY; PRT; 74 AA.
AC Q4SKC6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 13 SCAF14566, whole genome shotgun sequence.
DE (Fragment).
DE ORFNames=GSTENG00016792001;
GN Tetraodon nigroviridis (Green puffer).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014566; CAF98906.1; -; Genomic_DNA.
FT NON_TER 74 74
```

Search completed: May 12, 2006, 10:50:56
Job time : 106.769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 87.1795 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTLREWLHGGFC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	12	2	Aaw09524 Thrombopo
2	78	100.0	12	2	Aaw36675 Thrombopo
3	78	100.0	12	3	Aab16993 TPO-mimet
4	78	100.0	12	4	Aau25894 Human thr
5	78	100.0	12	5	Abb72879 TPO mimet
6	78	100.0	12	7	Adj73030 TPO mimet
7	78	100.0	12	8	Adj52665 CH1 delet
8	78	100.0	12	8	Adj51626 CH1 delet
9	78	100.0	16	2	Aaw09464 Thrombopo
10	78	100.0	16	2	Aaw33329 Thrombopo
11	78	100.0	16	3	Aab17019 TPO-mimet
12	78	100.0	16	4	Aau25829 Human thr
13	78	100.0	16	5	Abb72905 TPO mimet
14	78	100.0	16	7	Adj73057 TPO mimet
15	78	100.0	16	8	Adj52692 CH1 delet
16	78	100.0	16	8	Adj51653 CH1 delet
17	59	75.6	12	2	Aaw09582 Thrombopo
18	59	75.6	12	2	Aaw36733 Thrombopo
19	59	75.6	12	4	Aau25952 Human thr
20	54	69.2	12	2	Aaw09572 Thrombopo
21	54	69.2	12	2	Aaw36723 Thrombopo
22	54	69.2	12	3	Aab16994 TPO-mimet
23	54	69.2	12	4	Aau25942 Human thr
24	54	69.2	12	5	Abb72880 TPO mimet

98 43 55.1 10 5 ABB72886 Abb72886 TPO mimet 171
99 43 55.1 10 8 ADJ52672 Adj52672 CH1 delet 172
100 43 55.1 10 8 ADJ51633 Adj51633 CH1 delet 173
101 43 55.1 11 3 AAB17001 Aab17001 TPO-mimet 174
102 43 55.1 11 5 ABB72887 Abb72887 TPO mimet 175
103 43 55.1 11 8 ADJ52673 Adj52673 CH1 delet 176
104 43 55.1 11 8 ADJ51634 Adj51634 CH1 delet 177
105 43 55.1 13 3 AAB17003 Aab17003 TPO-mimet 178
106 43 55.1 13 5 ABB72889 Abb72889 TPO mimet 179
107 43 55.1 13 8 ADJ52675 Adj52675 CH1 delet 180
108 43 55.1 13 8 ADJ51636 Adj51636 CH1 delet 181
109 43 55.1 14 3 AAB17004 Aab17004 TPO-mimet 182
110 43 55.1 14 5 ABB72890 Abb72890 TPO mimet 183
111 43 55.1 14 8 ADJ52676 Adj52676 CH1 delet 184
112 43 55.1 14 8 ADJ51637 Adj51637 CH1 delet 185
113 43 55.1 18 2 AAW09583 Aaw09583 Thrombopo 186
114 43 55.1 18 2 AAW36734 Aaw36734 Thrombopo 187
115 43 55.1 18 4 AAU25953 Aau25953 Human thr 188
116 43 55.1 583 5 ABP65849 Abp65849 Bifidobac 189
117 43 55.1 791 6 AAE36091 Aae36091 Human dia 190
118 43 55.1 791 6 AAE36092 Aae36092 Human dia 191
119 43 55.1 791 7 ADE31701 Ade31701 Human 348 192
120 42 53.8 12 2 AAW09568 Aaw09568 Thrombopo 193
121 42 53.8 12 2 AAW09569 Aaw09569 Thrombopo 194
122 42 53.8 12 2 AAW36719 Aaw36719 Thrombopo 195
123 42 53.8 12 2 AAW36720 Aaw36720 Thrombopo 196
124 42 53.8 12 4 AAU25939 Aau25939 Human thr 197
125 42 53.8 12 4 AAU25938 Aau25938 Human thr 198
126 42 53.8 14 2 AAW09541 Aaw09541 Thrombopo 199
127 42 53.8 14 2 AAW36770 Aaw36770 Thrombopo 200
128 42 53.8 14 2 AAW36692 Aaw36692 Thrombopo 201
129 42 53.8 14 3 AAB16999 Aab16999 TPO-mimet 202
130 42 53.8 14 4 AAU25911 Aau25911 Human thr 203
131 42 53.8 14 5 ABB72885 Abb72885 TPO mimet 204
132 42 53.8 14 7 ADJ73036 Adj73036 TPO mimet 205
133 42 53.8 14 8 ADJ52671 Adj52671 CH1 delet 206
134 42 53.8 14 8 ADJ51632 Adj51632 CH1 delet 207
135 42 53.8 18 2 AAW09596 Aaw09596 Thrombopo 208
136 42 53.8 18 2 AAW36747 Aaw36747 Thrombopo 209
137 42 53.8 18 2 AAW36748 Aaw36748 Thrombopo 210
138 42 53.8 18 4 AAU25966 Aau25966 Human thr 211
139 42 53.8 18 4 AAU25967 Aau25967 Human thr 212
140 42 53.8 44 8 ABO59773 Abo59773 Human gen 213
141 42 53.8 56 4 ABG00428 Abg00428 Novel hum 214
142 42 53.8 173 9 ADY28796 Ady28796 CD47 bind 215
143 42 53.8 224 7 ABO71269 Abo71269 Pseudomon 216
144 42 53.8 287 9 AEB339606 Aeb339606 L. pneumo 217
145 42 53.8 287 9 AEB36183 Aeb36183 L. pneumo 218
146 42 53.8 1200 4 ABB63533 Abb63533 Drosophil 219
147 42 53.8 1481 4 ABG13697 Abg13697 Novel hum 220
148 41.5 257 8 ADS21782 Ads21782 Bacterial 221
149 41 52.6 10 2 AAW09547 Aaw09547 Thrombopo 222
150 41 52.6 10 2 AAW36698 Aaw36698 Thrombopo 223
151 41 52.6 10 4 AAU25917 Aau25917 Human thr 224
152 41 52.6 12 2 AAW09580 Aaw09580 Thrombopo 225
153 41 52.6 12 2 AAW36731 Aaw36731 Thrombopo 226
154 41 52.6 12 4 AAU25950 Aau25950 Human thr 227
155 41 52.6 218 8 ADR99369 Adr99369 Chlamydia 228
156 41 52.6 242 2 AAY35100 Aay35100 Chlamydia 229
157 41 52.6 306 7 ADM26482 Adm26482 Hyperther 230
158 41 52.6 310 8 ADT55980 Adt55980 Plant pol 231
159 41 52.6 486 6 ABP77210 Abp77210 N. gonorr 232
160 41 52.6 489 4 AAB92701 Aab92701 Human pro 233
161 41 52.6 513 2 AAY38789 Aay38789 Neisseria 234
162 41 52.6 513 2 AAY38786 Aay38786 Neisseria 235
163 41 52.6 513 2 AAY38787 Aay38787 Neisseria 236
164 41 52.6 513 2 AAY38788 Aay38788 Neisseria 237
165 41 52.6 513 9 AEB49392 Aeb49392 N. mening 238
166 41 52.6 513 9 AEB49394 Aeb49394 N. mening 239
167 41 52.6 513 9 AEB49395 Aeb49395 N. gonorr 240
168 41 52.6 513 9 AEB49397 Aeb49397 N. gonorr 241
169 41 52.6 604 6 ADA48262 Ada48262 Rice prot 242
170 41 52.6 604 7 ABM86602 Abm86602 Rice abio 243

41 52.6 619 4 AAU33715 Aau33715 Pseudomon
41 52.6 619 6 ABU15512 Abul5512 Protein e
41 52.6 770 3 AAY54060 Aay54060 Amino aci
41 52.6 776 3 AAY54062 Aay54062 Amino aci
41 52.6 804 7 ABO75043 Abo75043 Pseudomon
41 52.6 826 4 ABG21979 Abg21979 Novel hum
41 52.6 833 3 AAY54059 Aay54059 Amino aci
41 52.6 839 3 AAY54061 Aay54061 Amino aci
41 52.6 853 3 AAY54058 Aay54058 Amino aci
41 52.6 853 8 ADU06633 Adu06633 Novel bro
41 52.6 859 3 AAY54056 Aay54056 Amino aci
41 52.6 1025 5 ABB90967 Abb90967 Herbicida
41 52.6 1025 8 ADR04093 Adr04093 A thalian
41 52.6 1043 6 ABU15501 Abul5501 Protein e
41 52.6 1151 7 ABO74850 Abo74850 Pseudomon
41 52.6 4342 4 AAU33611 Aau33611 Pseudomon
41 52.6 4342 6 ABU15500 Abul5500 Protein e
40.5 51.9 3588 8 ADQ39441 Adq39441 Human myo
40.5 51.9 4346 8 ADQ39440 Adq39440 Human myo
40.5 51.9 4347 8 ADQ39439 Adq39439 Human myo
40.5 51.9 4370 7 ADJ69461 Adj69461 Human hea
40.5 51.9 4391 6 AAE34390 Aae34390 Human per
40.5 51.9 4393 4 AAB31889 Aab31889 Amino aci
40.5 51.9 4393 8 ADL35758 Adl35758 Human per
40.5 51.9 4393 8 ADQ39442 Adq39442 Human myo
40.5 51.9 4436 4 ABG23265 Abg23265 Novel hum
12 2 AAW09525 Aaw09525 Thrombopo
12 2 AAW09529 Aaw09529 Thrombopo
12 2 AAW09528 Aaw09528 Thrombopo
12 2 AAW09567 Aaw09567 Thrombopo
12 2 AAW09526 Aaw09526 Thrombopo
12 2 AAW36676 Aaw36676 Thrombopo
12 2 AAW36679 Aaw36679 Thrombopo
12 2 AAW36677 Aaw36677 Thrombopo
12 2 AAW36718 Aaw36718 Thrombopo
12 2 AAW36680 Aaw36680 Thrombopo
12 4 AAU25895 Aau25895 Human thr
12 4 AAU25937 Aau25937 Human thr
12 4 AAU25896 Aau25896 Human thr
12 4 AAU25898 Aau25898 Human thr
12 4 AAU25899 Aau25899 Human thr
13 3 AAB16995 Aab16995 TPO-mimet
13 5 ABB72881 Abb72881 TPO mimet
13 7 ADJ73032 Adj73032 TPO mimet
13 8 ADJ52667 Adj52667 CH1 delet
13 8 ADJ51628 Adj51628 CH1 delet
14 2 AAW09536 Aaw09536 Thrombopo
14 2 AAW36687 Aaw36687 Thrombopo
14 2 AAW36768 Aaw36768 Thrombopo
14 4 AAU25906 Aau25906 Human thr
18 2 AAW09592 Aaw09592 Thrombopo
18 2 AAW09595 Aaw09595 Thrombopo
18 2 AAW36746 Aaw36746 Thrombopo
18 2 AAW36743 Aaw36743 Thrombopo
18 4 AAU25965 Aau25965 Human thr
18 4 AAU25962 Aau25962 Human thr
66 3 AAG53141 Aag53141 Arabidops
67 7 ADC27609 Adc27609 Human col
73 4 ABG00429 Abg00429 Novel hum
77 8 ADX68706 Adx68706 Plant ful
80 4 AAU17763 Aau17763 Novel hum
80 7 ADG41143 Adg41143 Human res
80 7 ADI96917 Adi96917 Human res
140 8 ADP80867 Adp80867 Mouse par
169 7 ABM86829 Abm86829 Rice abio
176 5 AAE25747 Aae25747 Rice Ft h
176 8 ADR04263 Adr04263 Rice Ft h
231 6 ABU94266 Abu94266 Lolium pe
286 3 AAG06809 Aag06809 Arabidops
364 8 ADX89766 Adx89766 Plant ful
380 7 ADE56536 Ade56536 Rat Prote
408 6 ABU33917 Abu33917 Protein e
480 3 AAG06808 Aag06808 Arabidops

244	40	51.3	483	9	ABM93529	Abm93529 M. xanthu	317	39	50.0	1472	7	ADE57585	Ade57585 Rat Prote
245	40	51.3	510	3	AAG06807	Aag06807 Arabidops	318	39	50.0	1472	7	ADD47124	Add47124 Rat Prote
246	40	51.3	775	8	ADX93666	Adx93666 Plant ful	319	39	50.0	1472	7	ADE63422	Ade63422 Rat Prote
247	40	51.3	1567	4	ABB64992	Abb64992 Drosophil	320	39	50.0	1472	7	ADE83343	Ade83343 Rat Prote
248	40	51.3	1872	2	AAW68510	Aaw68510 Partial h	321	39	50.0	1472	9	AEB22116	Aeb22116 Rat alpha
249	40	51.3	2000	8	ADP23496	Adp23496 PRO polyp	322	39	50.0	1474	2	AAR11334	Aar11334 Recombina
250	40	51.3	2180	8	ADN22356	Adn22356 Bacterial	323	39	50.0	1474	3	AA97157	Aay97157 Human alp
251	40	51.3	2318	8	ADJ76296	Adj76296 Marker ge	324	39	50.0	1474	4	AAB50673	Aab50673 Human alp
252	40	51.3	2321	2	AAW49698	Aaw49698 Human Not	325	39	50.0	1474	5	AAU81017	Aau81017 Human alp
253	40	51.3	2321	7	ADJ69828	Adj69828 Human hea	326	39	50.0	1474	5	AAU74798	Aau74798 Human alp
254	40	51.3	2321	8	ADH34621	Adh34621 Notch hom	327	39	50.0	1474	5	ABP65218	Abp65218 Hypoxia-r
255	40	51.3	2321	8	ADJ75570	Adj75570 Marker ge	328	39	50.0	1474	7	AAO31219	Aao31219 Human alp
256	40	51.3	2321	8	ADN48701	Adn48701 Human Not	329	39	50.0	1474	7	AAO31224	Aao31224 Human alp
257	40	51.3	2321	9	ADX70405	Adx70405 Human Not	330	39	50.0	1474	7	AAO31221	Aao31221 Human alp
258	39	50.0	16	2	AAW09609	Aaw09609 Thrombopo	331	39	50.0	1474	7	AAO31222	Aao31222 Human alp
259	39	50.0	16	2	AAW36760	Aaw36760 Thrombopo	332	39	50.0	1474	7	AAO31224	Aao31224 Human alp
260	39	50.0	16	4	AAU25979	Aau25979 Human thr	333	39	50.0	1474	7	ADE63424	Ade63424 Human Pro
261	39	50.0	51	8	ADS52159	Ads52159 Human sec	334	39	50.0	1474	7	ADE57586	Ade57586 Human Pro
262	39	50.0	58	8	ADS52157	Ads52157 Human sec	335	39	50.0	1474	7	ADD47130	Add47130 Human Pro
263	39	50.0	67	4	AAW16701	Aam16701 Peptide #	336	39	50.0	1474	7	ADE60381	Ade60381 Human Pro
264	39	50.0	67	4	ABB35684	Abb35684 Peptide #	337	39	50.0	1474	7	ADD47126	Add47126 Human Pro
265	39	50.0	67	4	AAW29188	Aam29188 Peptide #	338	39	50.0	1474	7	ADE83344	Ade83344 Human Pro
266	39	50.0	67	4	ABB21115	Abb21115 Protein #	339	39	50.0	1474	7	ADD89034	Add89034 TAT275. 1
267	39	50.0	67	4	AAW68879	Aam68879 Human bon	340	39	50.0	1474	8	ADK41559	Adk41559 Anti-cell
268	39	50.0	67	4	AAW56498	Aam56498 Human bra	341	39	50.0	1474	8	ADP24808	Adp24808 PRO polyp
269	39	50.0	67	4	ABG50540	Abg50540 Human liv	342	39	50.0	1474	9	ADQ39717	Adq39717 Human myo
270	39	50.0	67	5	ABG38458	Abg38458 Human pep	343	39	50.0	1479	5	ADW96863	Adw96863 Human nas
271	39	50.0	96	5	ABP04289	Abp04289 Human ORF	344	39	50.0	1479	8	ABO10075	Ado10075 Novel hum
272	39	50.0	113	5	ABB08067	Abb08067 Human alp	345	39	50.0	1482	7	ADD47134	Add47134 Human Pro
273	39	50.0	174	9	ADY28813	Ady28813 CD4 bindi	346	39	50.0	1482	7	ADF90915	Adf90915 Human hep
274	39	50.0	183	5	ABB08068	Aay08067 Rat serin	347	39	50.0	1482	7	AAR11749	Aar11749 Human alp
275	39	50.0	186	2	AA9708067	Adn24475 Bacterial	348	39	50.0	1484	2	ABG28006	Abg28006 Novel hum
276	39	50.0	190	8	ADN24475	Abp05718 Human ORF	349	39	50.0	1487	4	ABG28006	Abg28006 Novel hum
277	39	50.0	192	5	ABP05718	Adn21721 Bacterial	350	39	50.0	1500	7	AAO31218	Aao31218 Human alp
278	39	50.0	201	8	ADN21721	Abu20005 Protein e	351	39	50.0	1542	4	ABG28005	Abg28005 Novel hum
279	39	50.0	206	6	ABU20005	Abu21268 Protein e	352	39	50.0	1542	7	ADE09990	Ade09990 Novel pro
280	39	50.0	206	6	ABU21268	Abo78209 Pseudomon	353	39	50.0	1543	7	ADE08651	Ade08651 Novel pro
281	39	50.0	218	7	ABO78209	Abu39399 Protein e	354	39	50.0	2531	5	ADI16935	Adi16935 Rat NOVX
282	39	50.0	254	6	ABU39399	Aaw41989 Flea seri	355	39	50.0	2531	5	ADI16934	Adi16934 Rat NOVX
283	39	50.0	260	2	AAW41989	Aab50622 Flea seri	356	39	50.0	2531	7	ADE63713	Ade63713 Rat Prote
284	39	50.0	260	4	AAW50622	Aaw34527 hTCP prot	357	39	50.0	2531	7	ADE63705	Ade63705 Rat Prote
285	39	50.0	268	2	AAW34527	Aaw73931 T-lymphoc	358	39	50.0	2531	7	ADE63709	Ade63709 Rat Prote
286	39	50.0	288	2	AAW73931	Aay18077 CD4+ T-ly	359	39	50.0	2531	7	ADE63701	Ade63701 Rat Prote
287	39	50.0	288	2	AA9718077	Aag31478 Arabidops	360	39	50.0	2531	8	ADO58587	Ado58587 Mouse Not
288	39	50.0	307	3	AAG31478	Ady86672 Sbfl rest	361	39	50.0	2531	9	ADZ12495	Adz12495 Murine Ca
289	39	50.0	323	9	ADY86687	Ady86687 Sbfl rest	362	38.5	49.4	41	7	ADN59772	Adn59772 Peptide-v
290	39	50.0	323	9	ADY866742	Ady53330 Streptomy	363	38.5	49.4	41	7	ADN59790	Adn59790 Peptide-v
291	39	50.0	339	9	ADY53330	Aag31477 Arabidops	364	38.5	49.4	46	7	ADN59784	Adn59784 Peptide-v
292	39	50.0	358	3	AAG31477	Ade09375 Novel pro	365	38.5	49.4	46	7	AAW09578	Aaw09578 Thrombopo
293	39	50.0	375	7	ADE09375	Abr53688 Protein s	366	38	48.7	12	2	AAW09578	Aaw09578 Thrombopo
294	39	50.0	384	6	ABR53688	Adk64056 Disease t	367	38	48.7	12	2	AAW36729	Aaw36729 Thrombopo
295	39	50.0	384	7	ADK64056	Adx92629 Plant ful	368	38	48.7	12	4	AAU25948	Aau25948 Human thr
296	39	50.0	450	8	ADX92629	Aag31476 Arabidops	369	38	48.7	18	2	AAW09585	Aaw09585 Thrombopo
297	39	50.0	492	3	AAG31476	Abb92817 Herbicida	370	38	48.7	18	2	AAW09591	Aaw09591 Thrombopo
298	39	50.0	492	5	ABB92817	Aaw73934 T-lymphoc	371	38	48.7	18	2	AAW09586	Aaw09586 Thrombopo
299	39	50.0	799	2	AAW73934	Aay18073 CD4+ T-ly	372	38	48.7	18	2	AAW36737	Aaw36737 Thrombopo
300	39	50.0	799	2	AA9718073	Abm82746 Human dia	373	38	48.7	18	2	AAW36736	Aaw36736 Thrombopo
301	39	50.0	841	8	ABM82746	Abul1716 Human MDD	374	38	48.7	18	4	AAU25956	Aau25956 Human thr
302	39	50.0	842	6	ABU11716	Aaw73933 T-lymphoc	375	38	48.7	18	4	AAU25955	Aau25955 Human thr
303	39	50.0	881	2	AAW73933	Aay18072 CD4+ T-ly	376	38	48.7	18	4	AAU25961	Aau25961 Human thr
304	39	50.0	881	2	AA9718072	Adp65188 Human pro	377	38	48.7	19	4	AAU25961	Aau25961 Human thr
305	39	50.0	913	7	ADP65188	Adu64366 Human pro	378	38	48.7	37	3	AAB16504	Aab16504 Peptide w
306	39	50.0	913	8	ADU64366	Aaw52306 Human pro	379	38	48.7	37	3	AAB16590	Aab16590 Bacteriop
307	39	50.0	915	2	AAW52306	Aaw73932 T-lymphoc	380	38	48.7	45	4	AAW92581	Aam92581 Human dig
308	39	50.0	915	2	AAW73932	Aay18071 CD4+ T-ly	381	38	48.7	45	4	AAU22607	Aau22607 Novel hum
309	39	50.0	915	2	AA9718071	Abk43949 Human can	382	38	48.7	45	7	ADB32447	Adb32447 Human nov
310	39	50.0	1285	3	AAB43949	Adk41536 Anti-cell	383	38	48.7	55	4	AAB16504	Aae06044 Human gen
311	39	50.0	1450	7	ADK41536	Adw96864 Human mat	384	38	48.7	55	5	ABG33866	Abg33866 Human sec
312	39	50.0	1450	9	ADW96864	Aau81018 Human alp	385	38	48.7	55	6	ADA40593	Ada40593 Human sec
313	39	50.0	1451	5	AAU81018	Adk41537 Anti-cell	386	38	48.7	55	6	ABR47648	Abr47648 Human sec
314	39	50.0	1451	7	ADK41537	Add47128 Rat Prote	387	38	48.7	56	3	AA97067	Aay87067 Human sec
315	39	50.0	1472	7	ADD47128	Ade60379 Rat Prote	388	38	48.7	62	6	ABU45653	Abu45653 Protein e
316	39	50.0	1472	7	ADE60379		389	38	48.7	73	4	ABG26752	Abg26752 Novel hum

390 38 48.7 94 8 ADK47669 Adk47669 Streptoco
391 38 48.7 104 3 AAY87158 Aay87158 Human sec
392 38 48.7 104 4 AAE06135 Aae06135 Human gen
393 38 48.7 104 5 ABG33958 Abg33958 Human sec
394 38 48.7 241 9 AEB40561 Aeb40561 L. pneumo
395 38 48.7 246 8 ADX94643 Adx94643 Plant ful
396 38 48.7 250 9 AEB37237 Aeb37237 L. pneumo
397 38 48.7 257 5 ABG69047 Abg69047 Amino aci
398 38 48.7 257 8 ADN27256 Adn27256 Bacterial
399 38 48.7 284 4 ABG15780 Abg15780 Novel hum
400 38 48.7 300 4 AAM39907 Aam39907 Human pol
401 38 48.7 300 4 AAU08687 Aau08687 Human FCT
402 38 48.7 300 7 ADB32041 Adb32041 Human FCT
403 38 48.7 338 7 ADE56432 Ade56432 Rat Prote
404 38 48.7 338 8 ADQ20762 Adq20762 Human sof
405 38 48.7 339 2 AAW74896 Aaw74896 Human sec
406 38 48.7 339 5 ABG95348 Abg95348 Human nov
407 38 48.7 339 6 ABO34542 Abo34542 Region of
408 38 48.7 339 7 ADI23203 Adi23203 Novel hum
409 38 48.7 339 8 ADH74205 Adh74205 Human sec
410 38 48.7 355 8 ADS22232 Ads22232 Bacterial
411 38 48.7 360 3 AAG31508 Aag31508 Arabidops
412 38 48.7 370 6 ADA55399 Ada55399 Human pro
413 38 48.7 370 6 ABR41179 Abr41179 Human DIT
414 38 48.7 371 5 ABP69844 Abp69844 Human pol
415 38 48.7 372 8 ADP24070 Adp24070 PRO polyp
416 38 48.7 379 3 AAG47664 Aag47664 Arabidops
417 38 48.7 389 4 ABG13806 Abg13806 Novel hum
418 38 48.7 394 3 AAG31507 Aag31507 Arabidops
419 38 48.7 412 7 ABO69766 Abo69766 Pseudomon
420 38 48.7 413 3 AAG47663 Aag47663 Arabidops
421 38 48.7 421 8 ADS27311 Aab41816 Human ORF
422 38 48.7 431 3 AAB41816 Aab41816 Human ORF
423 38 48.7 432 8 ADS26952 Ads26952 Bacterial
424 38 48.7 432 8 ADS26578 Ads26578 Bacterial
425 38 48.7 445 5 ABB91957 Abb91957 Herbicida
426 38 48.7 461 3 AAG29255 Aag29255 Arabidops
427 38 48.7 482 3 AAG31506 Aag31506 Arabidops
428 38 48.7 482 5 ABB90925 Abb90925 Herbicida
429 38 48.7 494 4 AAU58735 Aau58735 Propionib
430 38 48.7 494 6 ABM55254 Abm55254 Propionib
431 38 48.7 500 2 AAY41738 Aay41738 Human PRO
432 38 48.7 500 3 AAB44294 Aab44294 Human PRO
433 38 48.7 500 3 AAB47999 Aab47999 Cocoa cri
434 38 48.7 500 4 AAU29058 Aau29058 Human PRO
435 38 48.7 500 5 ABG90983 Abg90983 Breast sp
436 38 48.7 500 6 ABU58434 Abu58434 Human PRO
437 38 48.7 500 6 ABU87982 Abu87982 Novel hum
438 38 48.7 500 6 ABU84297 Abu84297 Human sec
439 38 48.7 500 6 ABR66171 Abr66171 Human sec
440 38 48.7 500 6 ABR65561 Abr65561 Human sec
441 38 48.7 500 6 ABU99501 Abu99501 Human sec
442 38 48.7 500 6 ABU82740 Abu82740 Human PRO
443 38 48.7 500 6 ABU89861 Abu89861 Novel hum
444 38 48.7 500 6 ABR68110 Abr68110 Human sec
445 38 48.7 500 6 ABR56795 Abr56795 Human sec
446 38 48.7 500 6 ABU96163 Abu96163 Novel hum
447 38 48.7 500 6 ABU92594 Abu92594 Human sec
448 38 48.7 500 6 ABO08671 Abo08671 Human sec
449 38 48.7 500 6 ABO02723 Abo02723 Human sec
450 38 48.7 500 6 ABR74877 Abr74877 Human sec
451 38 48.7 500 6 ABR94639 Abr94639 Human sec
452 38 48.7 500 6 ABO25240 Abo25240 Novel hum
453 38 48.7 500 6 ABU85612 Abu85612 Human PRO
454 38 48.7 500 6 ABU98772 Abu98772 Novel hum
455 38 48.7 500 6 ABU97987 Abu97987 Novel hum
456 38 48.7 500 6 ABU91693 Abu91693 Novel hum
457 38 48.7 500 6 ABU72246 Abu72246 Novel hum
458 38 48.7 500 6 ABU89386 Abu89386 Human PRO
459 38 48.7 500 6 ABU86227 Abu86227 Human sec
460 38 48.7 500 6 ABU67440 Abu67440 Human sec
461 38 48.7 500 6 ABU80468 Abu80468 Human PRO
462 38 48.7 500 6 ABR99386 Abr99386 Human sec

38 48.7 38 48.7 500 6 ABR98776 Abr98776 Human sec
38 48.7 500 6 ABO16299 Abo16299 Human sec
38 48.7 500 6 ABR92199 Abr92199 Human sec
38 48.7 500 6 ABO18840 Abo18840 Human sec
38 48.7 500 6 ABR78261 Abr78261 Human sec
38 48.7 500 6 ABU84997 Abu84997 Novel hum
38 48.7 500 6 ABO00136 Abo00136 Novel hum
38 48.7 500 6 ABO11468 Abo11468 Human sec
38 48.7 500 6 ABO02113 Abo02113 Human sec
38 48.7 500 6 ABU88687 Abu88687 Novel hum
38 48.7 500 6 ABU83382 Abu83382 Human sec
38 48.7 500 6 ABO06183 Abo06183 Novel hum
38 48.7 500 6 ABR59219 Abr59219 Human sec
38 48.7 500 6 ABO09281 Abo09281 Human sec
38 48.7 500 6 ABO19145 Abo19145 Novel hum
38 48.7 500 6 ABO11163 Abo11163 Human sec
38 48.7 500 6 ABR66781 Abr66781 Human sec
38 48.7 500 6 ABO15994 Abo15994 Human sec
38 48.7 500 6 ABO13700 Abo13700 Human sec
38 48.7 500 6 ABO13700 Abo13700 Human sec
38 48.7 500 6 ABU84926 Abu84926 Human sec
38 48.7 500 6 ABU65603 Abu65603 Human sec
38 48.7 500 6 ABO07451 Abo07451 Human PRO
38 48.7 500 6 ABO03638 Abo03638 Human sec
38 48.7 500 6 ABR67086 Abr67086 Human sec
38 48.7 500 6 ABO15689 Abo15689 Human sec
38 48.7 500 6 ABU55970 Abu55970 Human sec
38 48.7 500 6 ABU61124 Abu61124 Human PRO
38 48.7 500 6 ABU65298 Abu65298 Human PRO
38 48.7 500 6 ABU95243 Abu95243 Novel hum
38 48.7 500 6 ABU71146 Abu71146 Human PRO
38 48.7 500 6 ABO07756 Abo07756 Human PRO
38 48.7 500 6 ABR69997 Abr69997 Human sec
38 48.7 500 6 ABR69330 Abr69330 Human sec
38 48.7 500 6 ABO01471 Abo01471 Human PRO
38 48.7 500 6 ABU81273 Abu81273 Human PRO
38 48.7 500 6 ABR60070 Abr60070 Human sec
38 48.7 500 6 ABR67805 Abr67805 Human sec
38 48.7 500 6 ABR65193 Abr65193 Human sec
38 48.7 500 6 ABR68415 Abr68415 Human sec
38 48.7 500 6 ABR71827 Abr71827 Human sec
38 48.7 500 6 ABU85307 Abu85307 Human PRO
38 48.7 500 6 ABU88997 Abu88997 Human sec
38 48.7 500 6 ABU83077 Abu83077 Human sec
38 48.7 500 6 ABU94933 Abu94933 Novel hum
38 48.7 500 6 ABU90481 Abu90481 Novel hum
38 48.7 500 6 ABU83992 Abu83992 Human sec
38 48.7 500 6 ABU93643 Abu93643 Novel hum
38 48.7 500 6 ABR64888 Abr64888 Human sec
38 48.7 500 6 ABR68720 Abr68720 Human sec
38 48.7 500 6 ABO06536 Abo06536 Human sec
38 48.7 500 6 ABR99081 Abr99081 Human sec
38 48.7 500 6 ABU56965 Abu56965 Human PRO
38 48.7 500 6 ABU85917 Abu85917 Novel hum
38 48.7 500 6 ABU82204 Abu82204 Novel hum
38 48.7 500 6 ABU87215 Abu87215 Human PRO
38 48.7 500 6 ABU83687 Abu83687 Human sec
38 48.7 500 6 ABO08061 Abo08061 Human PRO
38 48.7 500 6 ABU81772 Abu81772 Novel hum
38 48.7 500 6 ABU65936 Abu65936 Novel hum
38 48.7 500 6 ABR59765 Abr59765 Human sec
38 48.7 500 6 ABU93953 Abu93953 Novel hum
38 48.7 500 6 ABU80393 Abu80393 Human sec
38 48.7 500 6 ABU99806 Abu99806 Novel hum
38 48.7 500 6 ABR66476 Abr66476 Human sec
38 48.7 500 6 ABR90894 Abr90894 Human sec
38 48.7 500 6 ABU94321 Abu94321 Human PRO
38 48.7 500 6 ABU79203 Abu79203 Human PRO
38 48.7 500 6 ABU86532 Abu86532 Human sec
38 48.7 500 6 ABU86837 Abu86837 Novel hum
38 48.7 500 6 ABU94626 Abu94626 Human PRO
38 48.7 500 6 ABO04553 Abo04553 Human PRO
38 48.7 500 6 ABR70302 Abr70302 Human sec
38 48.7 500 6 ABU98467 Abu98467 Human PRO

536	38	48.7	500	6	ABR65866	Abr65866 Human sec
537	38	48.7	500	6	ABR64583	Abr64583 Human sec
538	38	48.7	500	6	ABU79508	Abu79508 Human PRO
539	38	48.7	500	6	ABU92899	Abu92899 Human sec
540	38	48.7	500	6	ABU95858	Abu95858 Human PRO
541	38	48.7	500	6	ABU91078	Abu91078 Novel hum
542	38	48.7	500	6	ABU90171	Abu90171 Novel hum
543	38	48.7	500	6	ABO09586	AbO09586 Human sec
544	38	48.7	500	6	ABO10858	AbO10858 Human sec
545	38	48.7	500	6	ABR70912	Abr70912 Human sec
546	38	48.7	500	6	ABU87520	Abu87520 Human PRO
547	38	48.7	500	6	ABU91388	Abu91388 Human PRO
548	38	48.7	500	6	ABU84602	Abu84602 Human sec
549	38	48.7	500	6	ABR69692	Abr69692 Human sec
550	38	48.7	500	6	ABU80069	Abu80069 Human PRO
551	38	48.7	500	6	ABU93338	Abu93338 Human PRO
552	38	48.7	500	6	ABO09891	AbO09891 Human sec
553	38	48.7	500	6	ABO08976	AbO08976 Human sec
554	38	48.7	500	6	ABU10544	Abu10544 Human sec
555	38	48.7	500	6	ABU95553	Abu95553 Human PRO
556	38	48.7	500	6	ABU96762	Abu96762 Novel hum
557	38	48.7	500	6	ABR70607	Abr70607 Human sec
558	38	48.7	500	6	ABO04958	AbO04958 Novel hum
559	38	48.7	500	6	ABO08366	AbO08366 Human sec
560	38	48.7	500	6	ABO05573	AbO05573 Human sec
561	38	48.7	500	6	ABR73962	Abr73962 Human sec
562	38	48.7	500	6	ABR95554	Abr95554 Human sec
563	38	48.7	500	6	ABR80851	Abr80851 Human sec
564	38	48.7	500	6	ABR81156	Abr81156 Human sec
565	38	48.7	500	6	ABM00852	Abm00852 Human sec
566	38	48.7	500	6	ABR88454	Abr88454 Human sec
567	38	48.7	500	6	ABM77275	Abm77275 Human sec
568	38	48.7	500	6	ABO28759	AbO28759 Human sec
569	38	48.7	500	6	ABO31504	AbO31504 Human sec
570	38	48.7	500	6	ABM07921	Abm07921 Human sec
571	38	48.7	500	6	ABO40401	AbO40401 Human sec
572	38	48.7	500	6	ABO35826	AbO35826 Human PRO
573	38	48.7	500	6	ABO43965	AbO43965 Human PRO
574	38	48.7	500	6	ADA77822	Ada77822 Human sec
575	38	48.7	500	6	ABM24760	Abm24760 Human sec
576	38	48.7	500	6	ABO3028	AbO3028 Human sec
577	38	48.7	500	6	ABR90284	Abr90284 Human sec
578	38	48.7	500	6	ABM17198	Abm17198 Human sec
579	38	48.7	500	6	ABR94944	Abr94944 Human sec
580	38	48.7	500	6	ABR95249	Abr95249 Human sec
581	38	48.7	500	6	ABO21487	AbO21487 Human sec
582	38	48.7	500	6	ABR97751	Abr97751 Human sec
583	38	48.7	500	6	ABR87539	Abr87539 Human sec
584	38	48.7	500	6	ABM77580	Abm77580 Human sec
585	38	48.7	500	6	ABM27810	Abm27810 Human sec
586	38	48.7	500	6	ABM06091	Abm06091 Human sec
587	38	48.7	500	6	ABM03597	Abm03597 Human sec
588	38	48.7	500	6	ABM35048	Abm35048 Human sec
589	38	48.7	500	6	ABM26285	Abm26285 Human sec
590	38	48.7	500	6	ABO48067	AbO48067 Human sec
591	38	48.7	500	6	ABR92809	Abr92809 Human sec
592	38	48.7	500	6	ABO24570	AbO24570 Human sec
593	38	48.7	500	6	ABM11581	Abm11581 Human sec
594	38	48.7	500	6	ABM02682	Abm02682 Human sec
595	38	48.7	500	6	ABM15978	Abm15978 Human sec
596	38	48.7	500	6	ABO27539	AbO27539 Human sec
597	38	48.7	500	6	ABM29030	Abm29030 Human sec
598	38	48.7	500	6	ABM07006	Abm07006 Human sec
599	38	48.7	500	6	ABM21100	Abm21100 Human sec
600	38	48.7	500	6	ABM09446	Abm09446 Human sec
601	38	48.7	500	6	ABO41316	AbO41316 Human sec
602	38	48.7	500	6	ABO36131	AbO36131 Human PRO
603	38	48.7	500	6	ABO43660	AbO43660 Human PRO
604	38	48.7	500	6	ABM76360	Abm76360 Human sec
605	38	48.7	500	6	ABM76056	Abm76056 Human sec
606	38	48.7	500	6	ABM25675	Abm25675 Human sec
607	38	48.7	500	6	ABM25980	Abm25980 Human sec
608	38	48.7	500	6	ABO03333	AbO03333 Human sec
609	38	48.7	500	6	ABO2418	AbO2418 Human sec
610	38	48.7	500	6	ABR90589	Abr90589 Human sec
611	38	48.7	500	6	ABR73657	Abr73657 Human sec
612	38	48.7	500	6	ABO16909	AbO16909 Human sec
613	38	48.7	500	6	ABR94334	Abr94334 Human sec
614	38	48.7	500	6	ABR75841	Abr75841 Human sec
615	38	48.7	500	6	ABR71217	Abr71217 Human sec
616	38	48.7	500	6	ABR93114	Abr93114 Human sec
617	38	48.7	500	6	ABR93419	Abr93419 Human sec
618	38	48.7	500	6	ABR87844	Abr87844 Human sec
619	38	48.7	500	6	ABO27844	AbO27844 Human sec
620	38	48.7	500	6	ABO29979	AbO29979 Human sec
621	38	48.7	500	6	ABO33188	AbO33188 Human PRO
622	38	48.7	500	6	ABM04876	Abm04876 Human sec
623	38	48.7	500	6	ABM08836	Abm08836 Human sec
624	38	48.7	500	6	ABO36436	AbO36436 Human sec
625	38	48.7	500	6	ABO35521	AbO35521 Human PRO
626	38	48.7	500	6	ABO39486	AbO39486 Human sec
627	38	48.7	500	6	ABM10361	Abm10361 Human sec
628	38	48.7	500	6	ABM11886	Abm11886 Human sec
629	38	48.7	500	6	ABO52032	AbO52032 Human PRO
630	38	48.7	500	6	ABO52337	AbO52337 Human PRO
631	38	48.7	500	6	ABO23655	AbO23655 Human sec
632	38	48.7	500	6	ABR97141	Abr97141 Human sec
633	38	48.7	500	6	ABR86929	Abr86929 Human sec
634	38	48.7	500	6	ABM10971	Abm10971 Human sec
635	38	48.7	500	6	ABM28115	Abm28115 Human sec
636	38	48.7	500	6	ABO32114	AbO32114 Human sec
637	38	48.7	500	6	ABM15241	Abm15241 Human sec
638	38	48.7	500	6	ABM06396	Abm06396 Human sec
639	38	48.7	500	6	ABM04207	Abm04207 Human sec
640	38	48.7	500	6	ABM22320	Abm22320 Human sec
641	38	48.7	500	6	ABM07616	Abm07616 Human sec
642	38	48.7	500	6	ABO40706	AbO40706 Human sec
643	38	48.7	500	6	ABM35353	Abm35353 Human sec
644	38	48.7	500	6	ABM33116	Abm33116 Human sec
645	38	48.7	500	6	ABO52642	AbO52642 Human PRO
646	38	48.7	500	6	ABO50202	AbO50202 Human sec
647	38	48.7	500	6	ABU99196	Abu99196 Human sec
648	38	48.7	500	6	ABO04248	AbO04248 Human sec
649	38	48.7	500	6	ABO05878	AbO05878 Human sec
650	38	48.7	500	6	ABM18418	Abm18418 Human sec
651	38	48.7	500	6	ABR97446	Abr97446 Human sec
652	38	48.7	500	6	ABR80546	Abr80546 Human sec
653	38	48.7	500	6	ABM01157	Abm01157 Human sec
654	38	48.7	500	6	ABR88759	Abr88759 Human sec
655	38	48.7	500	6	ABM13411	Abm13411 Human sec
656	38	48.7	500	6	ABM20795	Abm20795 Human sec
657	38	48.7	500	6	ABO41926	AbO41926 Human sec
658	38	48.7	500	6	ABO42536	AbO42536 Human sec
659	38	48.7	500	6	ABM10056	Abm10056 Human sec
660	38	48.7	500	6	ABO38571	AbO38571 Human sec
661	38	48.7	500	6	ABM32811	Abm32811 Human sec
662	38	48.7	500	6	ABM22625	Abm22625 Human sec
663	38	48.7	500	6	ABM74836	Abm74836 Human sec
664	38	48.7	500	6	ADA79614	Ada79614 Human sec
665	38	48.7	500	6	ADA24902	Ada24902 Novel hum
666	38	48.7	500	6	ABR96226	Abr96226 Human sec
667	38	48.7	500	6	ABM02377	Abm02377 Human sec
668	38	48.7	500	6	ABR86319	Abr86319 Human sec
669	38	48.7	500	6	ABR86624	Abr86624 Human sec
670	38	48.7	500	6	ABM16588	Abm16588 Human sec
671	38	48.7	500	6	ABM29640	Abm29640 Human sec
672	38	48.7	500	6	ABO29064	AbO29064 Human sec
673	38	48.7	500	6	ABM23845	Abm23845 Human sec
674	38	48.7	500	6	ABM23235	Abm23235 Human sec
675	38	48.7	500	6	ABM22015	Abm22015 Human sec
676	38	48.7	500	6	ABO37656	AbO37656 Human sec
677	38	48.7	500	6	ABM28420	Abm28420 Human sec
678	38	48.7	500	6	ABM28725	Abm28725 Human sec
679	38	48.7	500	6	ABM66369	Abm66369 Human sec
680	38	48.7	500	6	ABM75751	Abm75751 Human sec
681	38	48.7	500	6	ABM34031	Abm34031 Human sec

682	38	48.7	500	6	ABM34336	Abm34336 Human sec	755	38	48.7	500	6	ABR94029	AbR94029 Human sec
683	38	48.7	500	6	ABO19695	AbO19695 Novel hum	756	38	48.7	500	6	ABR79936	AbR79936 Human sec
684	38	48.7	500	6	ABO20267	AbO20267 Human sec	757	38	48.7	500	6	ABM11276	AbM11276 Human sec
685	38	48.7	500	6	ABO21182	AbO21182 Human sec	758	38	48.7	500	6	ABO32883	AbO32883 Human PRO
686	38	48.7	500	6	ABO22097	AbO22097 Human sec	759	38	48.7	500	6	ABO30589	AbO30589 Human sec
687	38	48.7	500	6	ADA12563	AdA12563 Human sec	760	38	48.7	500	6	ABO30894	AbO30894 Human sec
688	38	48.7	500	6	ABR96531	AbR96531 Human sec	761	38	48.7	500	6	ABM27200	AbM27200 Human sec
689	38	48.7	500	6	ABR85709	AbR85709 Human sec	762	38	48.7	500	6	ABM29945	AbM29945 Human sec
690	38	48.7	500	6	ABR99691	AbR99691 Human sec	763	38	48.7	500	6	ABM05481	AbM05481 Human sec
691	38	48.7	500	6	ABM00242	AbM00242 Human sec	764	38	48.7	500	6	ABM15546	AbM15546 Human sec
692	38	48.7	500	6	ABM00547	AbM00547 Human sec	765	38	48.7	500	6	ABM08531	AbM08531 Human sec
693	38	48.7	500	6	ABO29674	AbO29674 Human sec	766	38	48.7	500	6	ABO42231	AbO42231 Human sec
694	38	48.7	500	6	ABM23540	AbM23540 Human sec	767	38	48.7	500	6	ABO37961	AbO37961 Human sec
695	38	48.7	500	6	ABM29335	AbM29335 Human sec	768	38	48.7	500	6	ABO45871	AbO45871 Human PRO
696	38	48.7	500	6	ABO38266	AbO38266 Human sec	769	38	48.7	500	6	ABM66674	AbM66674 Human sec
697	38	48.7	500	6	ABO45566	AbO45566 Human PRO	770	38	48.7	500	6	ADB20182	AdB20182 Human sec
698	38	48.7	500	6	ABM20490	AbM20490 Human sec	771	38	48.7	500	6	ABM19575	AbM19575 Human sec
699	38	48.7	500	6	ADA81341	AdA81341 Human sec	772	38	48.7	500	6	ABO49287	AbO49287 Human sec
700	38	48.7	500	6	ABO16604	AbO16604 Human sec	773	38	48.7	500	6	ABO49592	AbO49592 Human sec
701	38	48.7	500	6	ABO18230	AbO18230 Human sec	774	38	48.7	500	6	ADA78434	AdA78434 Human sec
702	38	48.7	500	6	ABO22657	AbO22657 Human PRO	775	38	48.7	500	6	ABO19586	AbO19586 Novel hum
703	38	48.7	500	6	ABO22962	AbO22962 Human PRO	776	38	48.7	500	6	ABR88149	AbR88149 Human sec
704	38	48.7	500	6	ABR92504	AbR92504 Human sec	777	38	48.7	500	6	ABM26895	AbM26895 Human sec
705	38	48.7	500	6	ABR81461	AbR81461 Human sec	778	38	48.7	500	6	ABM03292	AbM03292 Human sec
706	38	48.7	500	6	ABM77885	AbM77885 Human sec	779	38	48.7	500	6	ABO39791	AbO39791 Human sec
707	38	48.7	500	6	ABR89674	AbR89674 Human sec	780	38	48.7	500	7	ABO49897	AbO49897 Human sec
708	38	48.7	500	6	ABM26590	Abm26590 Human sec	781	38	48.7	500	7	ABO50812	AbO50812 Human sec
709	38	48.7	500	6	ABM13716	Abm13716 Human sec	782	38	48.7	500	7	ABO05268	AbO05268 Human sec
710	38	48.7	500	6	ABO28454	AbO28454 Human sec	783	38	48.7	500	7	ABR74572	AbR74572 Human sec
711	38	48.7	500	6	ABO30284	AbO30284 Human sec	784	38	48.7	500	7	ABR77051	AbR77051 Human sec
712	38	48.7	500	6	ABM07311	Abm07311 Human sec	785	38	48.7	500	7	ABM17808	AbM17808 Human sec
713	38	48.7	500	6	ABM03902	Abm03902 Human sec	786	38	48.7	500	7	ABR95859	AbR95859 Human sec
714	38	48.7	500	6	ABO37046	AbO37046 Human sec	787	38	48.7	500	7	ABO21792	AbO21792 Human sec
715	38	48.7	500	6	ABO41621	AbO41621 Human sec	788	38	48.7	500	7	ABO19962	AbO19962 Human sec
716	38	48.7	500	6	ABO35216	AbO35216 Human PRO	789	38	48.7	500	7	ABO24265	AbO24265 Human sec
717	38	48.7	500	6	ABM25065	Abm25065 Human sec	790	38	48.7	500	7	ABR86014	AbR86014 Human sec
718	38	48.7	500	6	ABO47457	AbO47457 Human sec	791	38	48.7	500	7	ABM10666	AbM10666 Human sec
719	38	48.7	500	6	ABO47762	AbO47762 Human sec	792	38	48.7	500	7	ABM76665	AbM76665 Human sec
720	38	48.7	500	6	ABO48372	AbO48372 Human sec	793	38	48.7	500	7	ABR89369	AbR89369 Human sec
721	38	48.7	500	6	ABO51422	AbO51422 Human PRO	794	38	48.7	500	7	ABM12496	AbM12496 Human sec
722	38	48.7	500	6	ABO51727	AbO51727 Human sec	795	38	48.7	500	7	ABM05786	Abm05786 Human sec
723	38	48.7	500	6	ABO50507	AbO50507 Human sec	796	38	48.7	500	7	ABO34911	AbO34911 Human PRO
724	38	48.7	500	6	ABR79631	AbR79631 Human sec	797	38	48.7	500	7	ABM02987	AbM02987 Human sec
725	38	48.7	500	6	ABM16893	Abm16893 Human sec	798	38	48.7	500	7	ABM18965	AbM18965 Human sec
726	38	48.7	500	6	ABO17925	AbO17925 Human sec	799	38	48.7	500	7	ABM19270	Abm19270 Human sec
727	38	48.7	500	6	ABO20877	AbO20877 Human sec	800	38	48.7	500	7	ABO46481	AbO46481 Human PRO
728	38	48.7	500	6	ABR96836	AbR96836 Human sec	801	38	48.7	500	7	ABO48982	AbO48982 Human sec
729	38	48.7	500	6	ABM12191	Abm12191 Human sec	802	38	48.7	500	7	ABR69025	AbR69025 Human sec
730	38	48.7	500	6	ABM16283	Abm16283 Human sec	803	38	48.7	500	7	ABR89064	AbR89064 Human sec
731	38	48.7	500	6	ABM24150	Abm24150 Human sec	804	38	48.7	500	7	ABR72437	AbR72437 Human sec
732	38	48.7	500	6	ABM14631	Abm14631 Human sec	805	38	48.7	500	7	ABR74267	AbR74267 Human sec
733	38	48.7	500	6	ABM04512	Abm04512 Human sec	806	38	48.7	500	7	ABO18535	AbO18535 Human sec
734	38	48.7	500	6	ABM06701	Abm06701 Human sec	807	38	48.7	500	7	ABR80241	AbR80241 Human sec
735	38	48.7	500	6	ABM09141	Abm09141 Human sec	808	38	48.7	500	7	ABM01462	Abm01462 Human sec
736	38	48.7	500	6	ABO39181	AbO39181 Human sec	809	38	48.7	500	7	ABM02072	Abm02072 Human sec
737	38	48.7	500	6	ABM75446	Abm75446 Human sec	810	38	48.7	500	7	ABR87234	AbR87234 Human sec
738	38	48.7	500	6	ABM25370	Abm25370 Human sec	811	38	48.7	500	7	ABM12801	Abm12801 Human sec
739	38	48.7	500	6	ABM19880	Abm19880 Human sec	812	38	48.7	500	7	ABM30555	Abm30555 Human sec
740	38	48.7	500	6	ABO46786	AbO46786 Human PRO	813	38	48.7	500	7	ABM24455	Abm24455 Human sec
741	38	48.7	500	6	ABO47091	AbO47091 Human PRO	814	38	48.7	500	7	ABO29369	AbO29369 Human sec
742	38	48.7	500	6	ADA83139	AdA83139 Human sec	815	38	48.7	500	7	ABO31199	AbO31199 Human sec
743	38	48.7	500	6	ABR71522	AbR71522 Human sec	816	38	48.7	500	7	ABM14326	Abm14326 Human sec
744	38	48.7	500	6	ABR72132	AbR72132 Human sec	817	38	48.7	500	7	ABM09751	Abm09751 Human sec
745	38	48.7	500	6	ABR98471	AbR98471 Human sec	818	38	48.7	500	7	ABO38876	AbO38876 Human sec
746	38	48.7	500	6	ABO06841	AbO06841 Human sec	819	38	48.7	500	7	ABM34641	Abm34641 Human sec
747	38	48.7	500	6	ABR84794	AbR84794 Human sec	820	38	48.7	500	7	ABO51117	AbO51117 Human sec
748	38	48.7	500	6	ABR73352	AbR73352 Human sec	821	38	48.7	500	7	ABO03943	AbO03943 Human sec
749	38	48.7	500	6	ABR76446	AbR76446 Human sec	822	38	48.7	500	7	ABO10413	AbO10413 Human PRO
750	38	48.7	500	6	ABR73047	AbR73047 Human sec	823	38	48.7	500	7	ABR77656	AbR77656 Human sec
751	38	48.7	500	6	ABM18113	Abm18113 Human sec	824	38	48.7	500	7	ABR78866	AbR78866 Human sec
752	38	48.7	500	6	ABO20572	AbO20572 Human sec	825	38	48.7	500	7	ABO23960	AbO23960 Human sec
753	38	48.7	500	6	ABO25315	AbO25315 Human PRO	826	38	48.7	500	7	ABR93724	AbR93724 Human sec
754	38	48.7	500	6	ABO25620	AbO25620 Human PRO	827	38	48.7	500	7	ABM01767	Abm01767 Human sec

828	38	48.7	500	7	ABM78190	Abm78190 Human sec	901	38	48.7	500	8	ADE74819	Ade74819 Human sec
829	38	48.7	500	7	ABR89979	Abr89979 Human sec	902	38	48.7	500	8	ADE89798	Ade89798 Human sec
830	38	48.7	500	7	ABM27505	Abm27505 Human sec	903	38	48.7	500	8	ADF61438	Adf61438 Human sec
831	38	48.7	500	7	ABM13106	Abm13106 Human sec	904	38	48.7	500	8	ADF40130	Adf40130 Human sec
832	38	48.7	500	7	ABO31809	Abo31809 Human sec	905	38	48.7	500	8	ADF45926	Adf45926 Human sec
833	38	48.7	500	7	ABM14021	Abm14021 Human sec	906	38	48.7	500	8	ADF24322	Adf24322 Human sec
834	38	48.7	500	7	ABM08226	Abm08226 Human sec	907	38	48.7	500	8	ADF40754	Adf40754 Human sec
835	38	48.7	500	7	ABO40096	Abo40096 Human sec	908	38	48.7	500	8	ADF23698	Adf23698 Human sec
836	38	48.7	500	7	ABM74531	Abm74531 Human sec	909	38	48.7	500	8	ADF33681	Adf33681 Human sec
837	38	48.7	500	7	ABM33726	Abm33726 Human sec	910	38	48.7	500	8	ADF27148	Adf27148 Human sec
838	38	48.7	500	7	ABM20185	Abm20185 Human sec	911	38	48.7	500	8	ADF27784	Adf27784 Human sec
839	38	48.7	500	7	ABO48677	Abo48677 Human sec	912	38	48.7	500	8	ADF41378	Adf41378 Human sec
840	38	48.7	500	7	ABR72742	Abr72742 Human sec	913	38	48.7	500	8	ADF33057	Adf33057 Human sec
841	38	48.7	500	7	ABO15384	Abo15384 Human sec	914	38	48.7	500	8	ADF25423	Adf25423 Human sec
842	38	48.7	500	7	ABR85099	Abr85099 Human sec	915	38	48.7	500	8	ADF26524	Adf26524 Human sec
843	38	48.7	500	7	ABO15079	Abo15079 Human sec	916	38	48.7	500	8	ADF34313	Adf34313 Human sec
844	38	48.7	500	7	ABO17214	Abo17214 Human sec	917	38	48.7	500	8	ADF46550	Adf46550 Human sec
845	38	48.7	500	7	ABM17503	Abm17503 Human sec	918	38	48.7	500	8	ADF96032	Adf96032 Novel hum
846	38	48.7	500	7	ABR85404	Abr85404 Human sec	919	38	48.7	500	8	ADG04303	Adg04303 Novel hum
847	38	48.7	500	7	ABM76970	Abm76970 Human sec	920	38	48.7	500	8	ADG00463	Adg00463 Novel hum
848	38	48.7	500	7	ABO28149	Abo28149 Human sec	921	38	48.7	500	8	ADG82719	Adg82719 Human PRO
849	38	48.7	500	7	ABM22930	Abm22930 Human sec	922	38	48.7	500	8	ADH26000	Adh26000 Novel hum
850	38	48.7	500	7	ABM30250	Abm30250 Human sec	923	38	48.7	500	8	ADG50536	Adg50536 Human sec
851	38	48.7	500	7	ABM21710	Abm21710 Human sec	924	38	48.7	500	8	ADG49912	Adg49912 Human sec
852	38	48.7	500	7	ABM21405	Abm21405 Human sec	925	38	48.7	500	8	ADG51784	Adg51784 Human sec
853	38	48.7	500	7	ABM14936	Abm14936 Human sec	926	38	48.7	500	8	ADH32969	Adh32969 Human PRO
854	38	48.7	500	7	ABO41011	Abo41011 Human sec	927	38	48.7	500	8	ADG49288	Adg49288 Human sec
855	38	48.7	500	7	ABO37351	Abo37351 Human sec	928	38	48.7	500	8	ADG48664	Adg48664 Human sec
856	38	48.7	500	7	ABO37351	Abo37351 Human sec	929	38	48.7	500	8	ADG51160	Adg51160 Human sec
857	38	48.7	500	7	ABM75141	Abm75141 Human sec	930	38	48.7	500	8	ADG59104	Adg59104 Human sec
858	38	48.7	500	7	ABM33421	Abm33421 Human sec	931	38	48.7	500	8	ADG62560	Adg62560 Human sec
859	38	48.7	500	7	ABO36421	Abo36421 Human PRO	932	38	48.7	500	8	ADH25585	Adh25585 Human neu
860	38	48.7	500	7	ADA82505	Ada82505 Human sec	933	38	48.7	500	8	ADJ54708	Adj54708 Human PRO
861	38	48.7	500	7	ABM31775	Abm31775 Human sec	934	38	48.7	500	8	ADJ64479	Adj64479 Human PRO
862	38	48.7	500	7	ABM31165	Abm31165 Human sec	935	38	48.7	500	8	ADM31375	Adm31375 Novel hum
863	38	48.7	500	7	ADB73869	Adb73869 Human PRO	936	38	48.7	500	8	Adm17362	Adm17362 Human sec
864	38	48.7	500	7	ADB85813	Adb85813 Human sec	937	38	48.7	500	8	ADM36422	Adm36422 Novel hum
865	38	48.7	500	7	ABM32080	Abm32080 Human sec	938	38	48.7	500	8	ADM40227	Adm40227 Novel hum
866	38	48.7	500	7	ABM32385	Abm32385 Human sec	939	38	48.7	500	8	AdL07196	Adl07196 Human sec
867	38	48.7	500	7	ABM31470	Abm31470 Human sec	940	38	48.7	500	8	Adn37835	Adn37835 Novel hum
868	38	48.7	500	7	ABM30860	Abm30860 Human sec	941	38	48.7	500	8	AdS98111	Ads98111 Protein f
869	38	48.7	500	7	ADB76585	Adb76585 Human PRO	942	38	48.7	500	8	ADU50318	Adu50318 PRO541, S
870	38	48.7	500	7	ADC44011	Adc44011 Human sec	943	38	48.7	500	9	Adw49597	Adw49597 PRO541 pr
871	38	48.7	500	7	ADC61771	Adc61771 Human sec	944	38	48.7	500	9	Adz52258	Adz52258 Human sec
872	38	48.7	500	7	ADC63735	Adc63735 Human sec	945	38	48.7	501	3	Aag47662	Aag47662 Arabidops
873	38	48.7	500	7	ADC66835	Adc66835 Human sec	946	38	48.7	506	8	AdS44990	Ads44990 Bacterial
874	38	48.7	500	7	ADC68959	Adc68959 Human sec	947	38	48.7	522	4	AAM41693	Aam41693 Human pol
875	38	48.7	500	7	ADC63019	Adc63019 Human sec	948	38	48.7	522	8	AdS98613	Ads98613 Protein f
876	38	48.7	500	7	ADC68084	Adc68084 Human sec	949	38	48.7	555	6	ABR83634	Abr83634 Saccharom
877	38	48.7	500	7	ADC41404	Adc41404 Human sec	950	38	48.7	578	3	AAG29254	Aag29254 Arabidops
878	38	48.7	500	7	ADC67459	Adc67459 Human sec	951	38	48.7	604	4	AAU16970	Aau16970 Human nov
879	38	48.7	500	7	ADC62395	Adc62395 Human sec	952	38	48.7	631	3	AAG29253	Aag29253 Arabidops
880	38	48.7	500	7	ADC42028	Adc42028 Human sec	953	38	48.7	631	7	AdF23256	Adf23256 AT4g31670
881	38	48.7	500	7	ADD05543	Add05543 Human sec	954	38	48.7	631	8	ADT55851	Adt55851 Plant pol
882	38	48.7	500	7	ADE49397	Ade49397 Human sec	955	38	48.7	642	8	Adn17690	Adn17690 Bacterial
883	38	48.7	500	7	ADE35451	Ade35451 Human sec	956	38	48.7	642	8	AdS42900	Ads42900 Bacterial
884	38	48.7	500	7	ADE16565	Ade16565 Human sec	957	38	48.7	648	8	Adf06702	Adf06702 Bacterial
885	38	48.7	500	7	ADD73180	Add73180 Human sec	958	38	48.7	658	8	AdS45120	Ads45120 Bacterial
886	38	48.7	500	7	ADD72538	Add72538 Human sec	959	38	48.7	658	7	ABO65348	Abo65348 Klebsiell
887	38	48.7	500	7	ADE17189	Ade17189 Human sec	960	38	48.7	717	8	ADN23120	Adn23120 Bacterial
888	38	48.7	500	7	ADF47203	Adf47203 Human sec	961	38	48.7	1376	7	AAE39971	Aae39971 Human hea
889	38	48.7	500	7	ADG02538	Adg02538 Novel hum	962	38	48.7	1376	7	ADJ68965	Adj68965 Human hea
890	38	48.7	500	7	ADF95420	Adf95420 Novel hum	963	38	48.7	1406	6	ABO00826	Abo00826 Polypepti
891	38	48.7	500	7	ADG11245	Adg11245 Novel hum	964	38	48.7	1406	7	AAE39972	Aae39972 Human CRU
892	38	48.7	500	7	ADG12235	Adg12235 Novel hum	965	38	48.7	1406	2	ADH11378	Adh11378 Vertebrat
893	38	48.7	500	7	ADH08895	Adh08895 Human PRO	966	38	48.7	1699	2	ADH11378	Adh11378 Vertebrat
894	38	48.7	500	7	ADG52960	Adg52960 Human sec	967	38	48.7	1808	2	ADH11234	Adh11234 Vertebrat
895	38	48.7	500	7	ADG60280	Adg60280 Human sec	968	38	48.7	2319	2	ADH11228	Adh11228 Vertebrat
896	38	48.7	500	7	ADI61040	Adi61040 Human sec	969	38	48.7	2732	2	ADH11251	Adh11251 Plastid t
897	38	48.7	500	7	ADL32676	Adl32676 Novel hum	970	38	48.7	4970	9	AEBO0355	Aeb00355 Thrombopo
898	38	48.7	500	7	ADM30210	Adm30210 Novel hum	971	37.5	48.1	14	2	AAW09535	Aaw09535 Thrombopo
899	38	48.7	500	8	ADE48697	Ade48697 Human sec	972	37.5	48.1	14	2	AAW36686	Aaw36686 Thrombopo
900	38	48.7	500	8	ADE74207	Ade74207 Human. sec	973	37.5	48.1	14	2	AAW46891	Aaw46891 Cyclic pe

974 37.5 48.1 14 3 AAB16997 Aab16997 TPO-mimet
975 37.5 48.1 14 4 AAU25905 Aau25905 Human thr
976 37.5 48.1 14 5 ABB72883 Abb72883 TPO mimet
977 37.5 48.1 14 6 ABG71743 Abg71743 TPO recep
978 37.5 48.1 14 7 ADJ73034 Adj73034 TPO mimet
979 37.5 48.1 14 8 ADJ52669 Adj52669 CH1 delet
980 37.5 48.1 14 8 ADJ51630 Adj51630 CH1 delet
981 37 47.4 9 4 AAM22332 Aam22332 HIV pepti
982 37 47.4 9 4 AAM22645 Aam22645 HIV pepti
983 37 47.4 9 4 AAM22647 Aam22647 HIV pepti
984 37 47.4 9 4 AAM22646 Aam22646 HIV pepti
985 37 47.4 10 2 AAW09560 Aaw09560 Thrombopo
986 37 47.4 10 2 AAW36711 Aaw36711 Thrombopo
987 37 47.4 10 4 AAU25930 Aau25930 Human thr
988 37 47.4 12 2 AAW09574 Aaw09574 Thrombopo
989 37 47.4 12 2 AAW09579 Aaw09579 Thrombopo
990 37 47.4 12 2 AAW36725 Aaw36725 Thrombopo
991 37 47.4 12 2 AAW36730 Aaw36730 Thrombopo
992 37 47.4 12 4 AAU25949 Aau25949 Human thr
993 37 47.4 12 4 AAU25944 Aau25944 Human thr
994 37 47.4 15 8 ADN17047 Adn17047 Gi-alpha
995 37 47.4 17 2 AAW09607 Aaw09607 Thrombopo
996 37 47.4 17 2 AAW09600 Aaw09600 Thrombopo
997 37 47.4 17 2 AAW09610 Aaw09610 Thrombopo
998 37 47.4 17 2 AAW36761 Aaw36761 Thrombopo
999 37 47.4 17 2 AAW36758 Aaw36758 Thrombopo
1000 37 47.4 17 2 AAW36751 Aaw36751 Thrombopo

ALIGNMENTS

RESULT 1
AAW09524
ID AAW09524 standard; protein; 12 AA.

AC AAW09524;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.

XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.
XX WO9640189-A1.
XX
PD 19-DEC-1996.

XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 28; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db |||||
1 CTLREWLHGGFC 12

RESULT 2
AAW36675
ID AAW36675 standard; peptide; 12 AA.

XX AAW36675;

XX 27-FEB-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

OS Synthetic.

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 28; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines

XX Sequence 12 AA;

Query Match 100.0%; Score 78; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWLHGGFC 12

RESULT 3

AAB16993
ID AAB16993 standard; peptide; 12 AA.
XX
AC AAB16993;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:49.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.

Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

Claim 19; Page 212; 608pp; English.

The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

Sequence 12 AA;

Query Match 100.0%; Score 78; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWLHGGFC 12

RESULT 4

AAU25894
ID AAU25894 standard; peptide; 12 AA.
XX
AC AAU25894;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #80.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; Fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.

Activating thrombopoietin receptors in cells, used to treat thrombocytopenia and hematological disorders, comprises contacting cells with peptides and peptide mimetics attached to hydrophilic polymers.

Disclosure; Col 21; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin (TPO) and to develop other compounds that bind to and activate the TPO receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

Sequence 12 AA;

Query Match 100.0%; Score 78; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWLHGGFC 12

Db 1 CTLREWLHGGFC 12

RESULT 5

ABB72879

ID ABB72879 standard; peptide; 12 AA.

XX AC ABB72879;

XX DT 05-APR-2002 (first entry)

XX DE TPO mimetic peptide SEQ ID NO:49.

XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP; TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianaemic; anorectic; antiinfertility; haemostatic; dermatological; neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Fanconi's syndrome.

XX OS Homo sapiens.

OS Synthetic.

XX PN WO200183525-A2.

XX PD 08-NOV-2001.

XX PF 02-MAY-2001; 2001WO-US014310.

XX PR 03-MAY-2000; 2000US-00563286.

XX PA (AMGE-) AMGEN INC.

XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX DR WPI; 2002-130313/17.

XX PT Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX PS Claim 39; Page 43; 176pp; English.

XX CC The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | | | |
Db 1 CTLREWLHGGFC 12

RESULT 6

ADJ73030

ID ADJ73030 standard; peptide; 12 AA.

XX AC ADJ73030;

XX DT 06-MAY-2004 (first entry)

XX DE TPO mimetic peptide sequence SeqID 484.

XX KW mimetic; CDR mimetibody; gene therapy; transgenic; immune; cardiovascular; infectious; malignant; neurologic disease; anaemia; immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective; TPO.

XX OS Synthetic.

XX PN WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX DR WPI; 2003-804237/75.

XX PT New CDR mimetibody comprising a portion of a heavy or light chain variable region comprising human framework or ligand binding region, useful for preparing a composition for treating e.g., immune, cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 484; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific portions or variants thereof. Specifically, it refers to an antibody fragment where a protein has been inserted into, or replaces a portion of, one or more CDR regions, such that each CDR mimetibody comprises at least one portion of a heavy chain or light chain variable region, which itself comprises at least one human framework region and at least one ligand binding region (LBR). The present invention describes human mimetibodies, including modified immunoglobulins and cleavage products that can be useful in gene therapy and the generation of transgenic plants and animals. Furthermore, the CDR mimetibody is useful for preparing compositions for modulating, treating or reducing the symptoms of immune, cardiovascular, infectious, malignant and/ or neurologic diseases, as well as anaemia. Accordingly, they exhibit immunomodulator, cardiant, antimicrobial, cytostatic and neuroprotective activities. This peptide sequence is a TPO mimetic peptide sequence used to make a mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | | | |

Query Match 100.0%; Score 78; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWLHGGFC 12

RESULT 9
AAW09464
ID AAW09464 standard; protein; 16 AA.
XX
AC AAW09464;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding compound peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 23; Page 90; 106pp; English.
XX

The present sequence is a compound which binds to thrombopoietin (TPO) receptor (TR). It has a molecular weight of < 8000 Da, and a binding affinity to TR as expressed by an IC50 of no more than about 100 mum. The compound can be used for treating patients suffering from haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. The peptide may also be used to maintain the proliferation and growth of TPO-dependent cell lines and for use in biological research, for detecting TPO receptors on living cells. The present sequence is a specifically claimed example of a compound that is administered to a patient suffering from a disorder that is susceptible to treatment with a thrombopoietin agonist

Sequence 16 AA;
Query Match 100.0%; Score 78; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 3 CTLREWLHGGFC 14

RESULT 10
AAW33329
ID AAW33329 standard; peptide; 16 AA.
XX
AC AAW33329;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX

KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 14; Page 89; 106pp; English.
XX

The present peptide binds the thrombopoietin receptor (TR), has a molecular weight of less than 8000 Da and a TR binding affinity as expressed by an IC50 of no more than about 100 microm. It can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

Sequence 16 AA;
Query Match 100.0%; Score 78; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 3 CTLREWLHGGFC 14

RESULT 11
AAB17019
ID AAB17019 standard; peptide; 16 AA.
XX

AC AAB17019;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:75.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer; autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;

KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
XX WO200024782-A2.
XX
XX PD 04-MAY-2000.
XX
XX PF 25-OCT-1999; 99WO-US025044.
XX
XX PR 23-OCT-1998; 98US-0105371P.
XX
XX PR 22-OCT-1999; 99US-00428082.
XX
XX PA (AMGE-) AMGEN INC.
XX
XX PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX DR WPI; 2000-350702/30.
XX
XX PT Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
XX PS Claim 19; Page 220; 608pp; English.
XX
XX CC The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX P3, and P4 = are each independently sequences of pharmacologically active
XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX cells from the present invention can be used for producing pharmaceutical
XX compositions. The compositions are useful for treating cancer, asthma,
XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX a Fab domain) can provide a longer half-life or incorporate functions
XX such as Fc receptor binding, protein A binding, complement fixation, and
XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX AAB18003 represent nucleotide and amino acid sequences used in the
XX exemplification of the present invention
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 78; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | | | |
Db 3 CTLREWLHGGFC 14

RESULT 12
AAU25829
ID AAU25829 standard; peptide; 16 AA.
XX
XX AC AAU25829;
XX
XX DT 17-DEC-2001 (first entry)
XX
XX DE Human thrombopoietin receptor (TPO-R) activator peptide #15.
XX
XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;

KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.
XX
OS Homo sapiens.
XX
XX PN US6251864-B1.
XX
XX PD 26-JUN-2001.
XX
XX PF 01-MAR-2000; 2000US-00516704.
XX
XX PR 07-JUN-1995; 95US-00478128.
XX
XX PR 07-JUN-1995; 95US-00485301.
XX
XX PR 07-JUN-1996; 96WO-US009623.
XX
XX PR 15-AUG-1996; 96US-00699027.
XX
XX PA (GLAX) GLAXO GROUP LTD.
XX
XX PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
XX PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
XX PI Yin Q;
XX
XX DR WPI; 2001-564142/63.
XX
XX PT Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
XX with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX PS Disclosure; Col 69-70; 128pp; English.
XX
XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX of activating thrombopoietin receptors in cells comprise contacting the
XX cells with effective amounts of peptides and peptide mimetics attached to
XX hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX as that due to chemotherapy, radiation therapy or bone-marrow
XX transplantation and to prevent thrombocytopenia in patients at risk.The
XX sequences are used to treat and prevent haematological disorders
XX including thrombocytopenia and platelet disorders. They are used in vitro
XX as unique tools for understanding the biological role of thrombopoietin
XX (TPO) and to develop other compounds that bind to and activate the TPO
XX receptor. The peptides can be used to detect TPO receptors on living
XX cells and fixed cells, in biological fluids, in tissue homogenates, and
XX in purified or natural biological materials. They may also be used for in
XX situ staining, fluorescence-activated cell sorting, Western blotting and
XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX be used for in vitro expansion of megakaryocytes and their committed
XX progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 78; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | | | |
Db 3 CTLREWLHGGFC 14

RESULT 13
ABB72905
ID ABB72905 standard; peptide; 16 AA.
XX
XX AC ABB72905;
XX
XX DT 05-APR-2002 (first entry)
XX
XX DE TPO mimetic peptide SEQ ID NO:75.
XX
XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX Homo sapiens.
OS Synthetic.
OS WO200183525-A2.
XX
XX
XX
PD 08-NOV-2001.
XX
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
XX
PR 03-MAY-2000; 2000US-00563286.
XX
XX
PA (AMGE-) AMGEN INC.
XX
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
XX WPI; 2002-130313/17.
DR
XX
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
XX
PS Claim 39; Page 44; 176pp; English.
XX
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytosstatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 78; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db |||||
3 CTLREWLHGGFC 14

RESULT 14
ADJ73057
ID ADJ73057 standard; peptide; 16 AA.
XX
AC ADJ73057;
XX
DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 511.
DE
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
XX Synthetic.
XX WO2003084477-A2.
XX
XX PD 16-OCT-2003.
XX
XX PF 24-MAR-2003; 2003WO-US009139.
XX
XX PR 29-MAR-2002; 2002US-0368791P.
XX
XX PA (CENZ) CENTOCOR INC.
XX
XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
XX DR WPI; 2003-804237/75.
XX
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
XX PS Disclosure; SEQ ID NO 511; 97pp; English.
XX
XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 78; DB 7; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db |||||
3 CTLREWLHGGFC 14

RESULT 15
ADJ52692
ID ADJ52692 standard; peptide; 16 AA.
XX
AC ADJ52692;
XX
XX DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID511.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082870/08.
DR
XX
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
XX Claim 2; SEQ ID NO 511; 129pp; English.
PS
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 16 AA;
Query Match 100.0%; Score 78; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTLREWLHGGFC 12
Db |||||
3 CTLREWLHGGFC 14
RESULT 16
ADJ51653
ID ADJ51653 standard; peptide; 16 AA.
XX
AC ADJ51653;
XX
XX 06-MAY-2004 (first entry)
DT
XX CH1 deleted mimetibody-related peptide SeqID511.
DE
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
XX Unidentified.
OS Synthetic.
OS
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
PF
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
PR
XX (CENZ) CENTOCOR INC.
PA
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
PI
XX
XX WPI; 2004-082872/08.
DR
XX
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 15; SEQ ID NO 511; 123pp; English.
PS
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC immunomodulator, antiallergic, muscular-Gen, hepatotropic, haemostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 16 AA;
Query Match 100.0%; Score 78; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTLREWLHGGFC 12
Db |||||
3 CTLREWLHGGFC 14
RESULT 17
AAW09582
ID AAW09582 standard; protein; 12 AA.
XX
AC AAW09582;
XX
DT 10-SEP-1997 (first entry)
XX

DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 12 AA;

Query Match 75.6%; Score 59; DB 2; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLQEWLSGGLC 12

RESULT 18
AAW36733
ID AAW36733 standard; peptide; 12 AA.
XX
AC AAW36733;
XX
DT 09-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX

PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 12 AA;

Query Match 75.6%; Score 59; DB 2; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLQEWLSGGLC 12

RESULT 19
AAU25952
ID AAU25952 standard; peptide; 12 AA.
XX
AC AAU25952;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #138.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;

XX WPI; 2001-564142/63.

DR

XX

PT Activating thrombopoietin receptors in cells, used to treat

PT thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX

PS Disclosure; Col 22; 128pp; English.

XX

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent hematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

CC progenitors alone or in conjunction with additional cytokines

XX

SQ Sequence 12 AA;

Query Match 75.6%; Score 59; DB 4; Length 12;

Best Local Similarity 75.0%; Pred. No. 0.05;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12

Db 1 CTLQEWLSGGLC 12

RESULT 20

AAW09572

ID AAW09572 standard; protein; 12 AA.

XX

AC AAW09572;

XX

DT 10-SEP-1997 (first entry)

XX

DE Thrombopoietin receptor binding peptide.

XX

KW Haematology; thrombocytopenia; TPO; TR; proliferation;

KW bone marrow transfusion; chemotherapy; radiation therapy.

XX

OS Synthetic.

XX

PN WO9640189-A1.

XX

PD 19-DEC-1996.

XX

PF 05-JUN-1996; 96WO-US008998.

XX

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX

PA (GLAX) GLAXO GROUP LTD.

XX

PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX

DR WPI; 1997-051883/05.

XX

PT Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

XX

PS Disclosure; Page 29; 106pp; English.

XX

CC The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering

CC from haematological disorders and thrombocytopenia resulting from

CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide

CC may also be used to maintain the proliferation and growth of TPO-

CC dependent cell lines and for use in biological research, for detecting

CC TPO receptors on living cells

XX

SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 2; Length 12;

Best Local Similarity 66.7%; Pred. No. 0.28;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12

Db 1 CTLREWVFAGLC 12

RESULT 21

AAW36723

ID AAW36723 standard; peptide; 12 AA.

XX

AC AAW36723;

XX

DT 09-MAR-1998 (first entry)

XX

DE Thrombopoietin receptor binding peptide.

XX

KW Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX

OS Synthetic.

XX

PN WO9640750-A1.

XX

PD 19-DEC-1996.

XX

PF 07-JUN-1996; 96WO-US009623.

XX

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX

PA (GLAX) GLAXO GROUP LTD.

XX

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX

DR WPI; 1997-052226/05.

XX

PT Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of haematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX

PS Disclosure; Page 29; 106pp; English.

XX

CC The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX SQ Sequence 12 AA; Query Match 69.2%; Score 54; DB 2; Length 12; Best Local Similarity 66.7%; Pred. No. 0.28; Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWVFAGLC 12

RESULT 22
AAB16994
ID AAB16994 standard; peptide; 12 AA.
XX
AC AAB16994;
XX 31-OCT-2000 (first entry)
XX TPO-mimetic peptide sequence SEQ ID NO:50.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS WO200024782-A2.
PN 04-MAY-2000.
XX 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX (AMGE-) AMGEN INC.
PA Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.
PT Claim 19; Page 212; 608pp; English.
XX The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention
XX SQ Sequence 12 AA;

XX Query Match 69.2%; Score 54; DB 3; Length 12; Best Local Similarity 66.7%; Pred. No. 0.28; Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWVFAGLC 12

RESULT 23
AAU25942
ID AAU25942 standard; peptide; 12 AA.
XX
AC AAU25942;
XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #128.
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX Homo sapiens.
OS US6251864-B1.
PN 26-JUN-2001.
XX 01-MAR-2000; 2000US-00516704.
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX (GLAX) GLAXO GROUP LTD.
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
XX Activating thrombopoietin receptors in cells, used to treat thrombocytopenia and hematological disorders, comprises contacting cells with peptides and peptide mimetics attached to hydrophilic polymers.
PS Disclosure; Col 22; 128pp; English.
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin (TPO) and to develop other compounds that bind to and activate the TPO receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines
XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 4; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWVFAGLC 12

RESULT 24
ABB72880
ID ABB72880 standard; peptide; 12 AA.
XX
AC ABB72880;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:50.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 5; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 1 CTLREWVFAGLC 12

RESULT 25
ADJ73031.
ID ADJ73031 standard; peptide; 12 AA.
XX
AC ADJ73031;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 485.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX WPI; 2003-804237/75.
DR
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 485; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 7; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | : | |
Db 1 CTLREWVFAGLC 12

RESULT 26
ADJ52666
ID ADJ52666 standard; peptide; 12 AA.

XX AC ADJ52666;
XX DT 06-MAY-2004 (first entry)
XX DE CH1 deleted mimetibody-related peptide SeqID485.

XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX OS Unidentified.
OS Synthetic.

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

XX PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX PS Claim 2; SEQ ID NO 485; 129pp; English.

XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTLREWLHGGFC 12
| | | | | : | |
Db 1 CTLREWVFAGLC 12

RESULT 27
ADJ51627
ID ADJ51627 standard; peptide; 12 AA.

XX AC ADJ51627;
XX DT 06-MAY-2004 (first entry)
XX DE CH1 deleted mimetibody-related peptide SeqID485.

XX KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunologic disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX OS Unidentified.
OS Synthetic.

XX PN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082872/08.

XX PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX CC Claim 14; SEQ ID NO 485; 123pp; English.

XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gynaecological-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | : | |
Db 1 CTLREWVFAGLC 12

RESULT 28
AAW09581
ID AAW09581 standard; protein; 12 AA.

XX AAW09581;
XX 10-SEP-1997 (first entry)
XX Thrombopoietin receptor binding peptide.
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX Synthetic.
XX WO9640189-A1.
XX 19-DEC-1996.
XX 05-JUN-1996; 96WO-US008998.
XX 07-JUN-1995; 95US-00472371.
XX 07-JUN-1995; 95US-00473604.
XX 07-JUN-1995; 95US-00476168.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00484090.
XX 07-JUN-1995; 95US-00485301.
XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 29; 106pp; English.
XX The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells

XX SQ Sequence 12 AA;
Query Match 66.7%; Score 52; DB 2; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | : | |

Db 1 CTLREWLSYGTG 12

RESULT 29
AAU25951

ID AAU25951 standard; peptide; 12 AA.

XX AAU25951;
XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #137.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, 'Cwirla SE, Gates CM, Schatz PJ;
XX Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
XX Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 22; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines

XX SQ Sequence 12 AA;

Query Match 66.7%; Score 52; DB 4; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12

Db 1 CTLREWLSYGTC 12

RESULT 30
AAW36732
ID AAW36732 standard; peptide; 13 AA.
XX
AC AAW36732;
XX
DT 09-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-0522226/05.
PA Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 13 AA;

Query Match 66.7%; Score 52; DB 2; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 2 CTLREWLSYGTC 13

RESULT 31
AAW09566
ID AAW09566 standard; protein; 10 AA.
XX
AC AAW09566;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX

KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 10 AA;

Query Match 62.8%; Score 49; DB 2; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
Db 1 CTLREWLEG 9

RESULT 32
AAW36717
ID AAW36717 standard; peptide; 10 AA.
XX
AC AAW36717;
XX
DT 27-FEB-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 29; 106pp; English.
PS
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 10 AA;

Query Match 62.8%; Score 49; DB 2; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTLREWLHG 9
Db 1 CTLREWLEG 9

RESULT 33
AAU25936
ID AAU25936 standard; peptide; 10 AA.
XX
AC AAU25936;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #122.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
XX US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 22; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 10 AA;

Query Match 62.8%; Score 49; DB 4; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTLREWLHG 9
Db 1 CTLREWLEG 9

RESULT 34
AAB17002
ID AAB17002 standard; peptide; 12 AA.
XX
AC AAB17002;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:58.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CRLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 214; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 61.5%; Score 48; DB 3; Length 12;
Best Local Similarity 66.7%; Pred. NO. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWLXXXXC 12

RESULT 35
ABB72888
ID ABB72888 standard; peptide; 12 AA.

XX
AC ABB72888;

XX
DT 05-APR-2002 (first entry)

XX
DE TPO mimetic peptide SEQ ID NO:58.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW antianemic; anorectic; antinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

XX
PD 08-NOV-2001.

XX
PF 02-MAY-2001; 2001WO-US014310.

XX
PR 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

DR

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
CC anianaemic, anorectic, antinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 12 AA;

Query Match 61.5%; Score 48; DB 5; Length 12;
Best Local Similarity 66.7%; Pred. NO. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWLXXXXC 12

RESULT 36

ADJ52674

ID ADJ52674 standard; peptide; 12 AA.

XX
AC ADJ52674;

XX
DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID493.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

XX
FH Key Location/Qualifiers

FT Misc-difference 1..12

FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

XX WO2004002417-A2.

XX
PD 08-JAN-2004.

XX
PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
PA Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;
PI WPI; 2004-082870/08.
XX
DR New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
XX Claim 2; SEQ ID NO 493; 129pp; English.
PS
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
XX Sequence 12 AA;
SQ

Query Match 61.5%; Score 48; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWLXXXXC 12

RESULT 37
ADJ51635
ID ADJ51635 standard; peptide; 12 AA.
XX
AC ADJ51635;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID493.
XX
KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
XX Unidentified.
OS
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .12

FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
XX 08-JAN-2004.
PD
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
PI
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 14; SEQ ID NO 493; 123pp; English.
PS
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
XX Sequence 12 AA;
SQ

Query Match 61.5%; Score 48; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWLXXXXC 12

RESULT 38
AAW09570
ID AAW09570 standard; protein; 12 AA.
XX
XX AAW09570;
AC
XX 10-SEP-1997 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
DE
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
KW
XX Synthetic.
OS
XX WO9640189-A1.
PN

XX 19-DEC-1996.
PD
XX
XX
PF 05-JUN-1996; 96WO-US008998.
XX
XX 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-051883/05.
DR
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 29; 106pp; English.
PS
XX The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
XX Sequence 12 AA;
SQ
Query Match 60.3%; Score 47; DB 2; Length 12;
Best Local Similarity 58.3%; Pred. No. 3.2;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTLREWLHGGFC 12
Db ||||:| :|
1 CTLRQWLGDWC 12
RESULT 39
AAW36721
ID AAW36721 standard; peptide; 12 AA.
XX
XX AAW36721;
AC
XX 09-MAR-1998 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
DE
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
OS
XX WO9640750-A1.
PN
XX 19-DEC-1996.
PD
XX 07-JUN-1996; 96WO-US009623.
PF
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
DR
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 29; 106pp; English.
PS
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX Sequence 12 AA;
SQ
Query Match 60.3%; Score 47; DB 2; Length 12;
Best Local Similarity 58.3%; Pred. No. 3.2;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTLREWLHGGFC 12
Db ||||:| :|
1 CTLRQWLGDWC 12
RESULT 40
AAU25940
ID AAU25940 standard; peptide; 12 AA.
XX
XX AAU25940;
AC
XX 17-DEC-2001 (first entry)
DT
XX Human thrombopoietin receptor (TPO-R) activator peptide #126.
DE
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
XX Homo sapiens.
OS
XX US6251864-B1.
XX
XX 26-JUN-2001.
PD
XX
XX 01-MAR-2000; 2000US-00516704.
PF
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
XX WPI; 2001-564142/63.
DR
XX Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
PT
XX Disclosure; Col 22; 128pp; English.
PS

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 12 AA;

Query Match 60.3%; Score 47; DB 4; Length 12;
Best Local Similarity 58.3%; Pred. No. 3.2;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
||||:|:|
Db 1 CTLRQWLGDWC 12

RESULT 41
AAW09507
ID AAW09507 standard; protein; 10 AA.

XX AAW09507;

XX 10-SEP-1997 (first entry)

Thrombopoietin receptor binding peptide.

Haematology; thrombocytopenia; TPO; TR; proliferation;
bone marrow transfusion; chemotherapy; radiation therapy.

Synthetic.

WO9640189-A1.

19-DEC-1996.

05-JUN-1996; 96WO-US008998.

07-JUN-1995; 95US-00472371.

07-JUN-1995; 95US-00473604.

07-JUN-1995; 95US-00476168.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00484090.

07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
WPI; 1997-051883/05.
Thrombopoietin receptor-binding/activating peptide(s) and peptide
mimetic(s) - useful in treatment of haematological disorders, esp.
thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 27; 106pp; English.

The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
||||:|:|
Db 1 CTLRQWLQG 9

RESULT 42
AAW36658

ID AAW36658 standard; peptide; 10 AA.

XX AAW36658;

XX 11-MAR-1998 (first entry)

Thrombopoietin receptor binding peptide.

Thrombopoietin receptor; binding peptide; treatment; agonist;
haematological disorder; thrombocytopenia; chemotherapy;
radiation therapy; bone marrow transfusion; diagnosis;
signal transduction; receptor activation; cell culture.

Synthetic.

WO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the
thrombopoietin receptor - useful in treatment of haematological
disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 27; 106pp; English.

The present peptide, which binds the thrombopoietin receptor (TR), can be
used to treat disorders which are susceptible to treatment with a
thrombopoietin agonist, preferably haematological disorders and
thrombocytopenia resulting from chemotherapy, radiation therapy or bone
marrow transfusions. It can also be used diagnostically, e.g. to
investigate the mechanism of thrombopoietin signal transduction and
receptor activation, or to maintain the proliferation and growth of
thrombopoietin dependent cell lines

Sequence 10 AA;

Query Match 59.0%; Score 46; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9

Db	:	1	CTLRQWLQG	9	
RESULT 43					
AAB16990					
ID	AAB16990	standard;	peptide;	10	AA.
XX					
AC	AAB16990;				
XX					
DT	31-OCT-2000	(first entry)			
XX					
DE	TPO-mimetic peptide sequence SEQ ID NO:46.				
XX					
KW	Modified peptide; therapeutic agent; fusion; Fc domain; cancer;				
KW	autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;				
KW	immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;				
KW	inhibitor; erythropoietin; thrombopoietin; interleukin 1;				
KW	cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;				
KW	vascular endothelial growth factor; matrix metalloproteinase; asthma;				
KW	thrombosis; pharmaceutical.				
XX					
OS	Synthetic.				
XX					
PN	WO200024782-A2.				
XX					
PD	04-MAY-2000.				
XX					
PF	25-OCT-1999;	99WO-US025044.			
XX					
PR	23-OCT-1998;	98US-0105371P.			
PR	22-OCT-1999;	99US-00428082.			
XX					
PA	(AMGE-) AMGEN INC.				
XX					
PI	Feige U, Liu C, Cheetham J, Boone TC;				
XX					
DR	WPI; 2000-350702/30.				
XX					
PT	Novel composition of matter comprising an Fc domain and pharmacologically				
PT	active peptides, useful for treating cancer and autoimmune diseases.				
XX					
PS	Claim 19; Page 211; 608pp; English.				
XX					
CC	The present invention describes composition of matter (I) comprising an				
CC	Fc domain, pharmacologically active peptides, and linkers. Where (I) is:				
CC	(X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each				
CC	independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-				
CC	(L2)d-P2-(L3)e-P ³ , or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,				
CC	P3, and P4 = are each independently sequences of pharmacologically active				
CC	peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,				
CC	c, d, e, and f = are each independently 0 or 1, provided that at least 1				
CC	of a and b is 1. The composition can have cytostatic, antiasthmatic,				
CC	thrombolytic and immunosuppressive activities. DNAs, vectors and host				
CC	cells from the present invention can be used for producing pharmaceutical				
CC	compositions. The compositions are useful for treating cancer, asthma,				
CC	thrombosis, or autoimmune diseases. The use of an Fc domain (rather than				
CC	a Fab domain) can provide a longer half-life or incorporate functions				
CC	such as Fc receptor binding, protein A binding, complement fixation, and				
CC	possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to				
CC	AAB18003 represent nucleotide and amino acid sequences used in the				
CC	exemplification of the present invention				
XX					
SQ	Sequence 10 AA;				
Query Match	59.0%;	Score 46;	DB 3;	Length 10;	
Best Local Similarity	77.8%;	Pred. No. 3.8;			
Matches	7;	Conservative	1;	Mismatches	1; Indels 0; Gaps 0;
QY	1	CTLRWLHG	9		
	:				
Db	1	CTLRQWLQG	9		
RESULT 44					
AAU25877					
ID	AAU25877	standard;	peptide;	10	AA.
XX					
AC	AAU25877;				
XX					
DT	17-DEC-2001	(first entry)			
XX					
DE	Human thrombopoietin receptor (TPO-R) activator peptide #63.				
XX					
KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;				
KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;				
KW	bone marrow transplantation; haematological disorder; platelet disorder;				
KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;				
KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;				
KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.				
OS	Homo sapiens.				
XX					
PN	US6251864-B1.				
XX					
PD	26-JUN-2001.				
XX					
PF	01-MAR-2000;	2000US-00516704.			
XX					
PR	07-JUN-1995;	95US-00478128.			
PR	07-JUN-1995;	95US-00485301.			
PR	07-JUN-1996;	96WO-US009623.			
PR	15-AUG-1996;	96US-00699027.			
XX					
PA	(GLAX) GLAXO GROUP LTD.				
XX					
PI	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;				
PI	Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;				
PI	Yin Q;				
XX					
DR	WPI; 2001-564142/63.				
XX					
PT	Activating thrombopoietin receptors in cells, used to treat				
PT	thrombocytopenia and hematological disorders, comprises contacting cells				
PT	with peptides and peptide mimetics attached to hydrophilic polymers.				
XX					
PS	Disclosure; Col 20; 128pp; English.				
XX					
CC	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that				
CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods				
CC	of activating thrombopoietin receptors in cells comprise contacting the				
CC	cells with effective amounts of peptides and peptide mimetics attached to				
CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such				
CC	as that due to chemotherapy, radiation therapy or bone-marrow				
CC	transplantation and to prevent thrombocytopenia in patients at risk. The				
CC	sequences are used to treat and prevent haematological disorders				
CC	including thrombocytopenia and platelet disorders. They are used in vitro				
CC	as unique tools for understanding the biological role of thrombopoietin				
CC	(TPO) and to develop other compounds that bind to and activate the TPO				
CC	receptor. The peptides can be used to detect TPO receptors on living				
CC	cells and fixed cells, in biological fluids, in tissue homogenates, and				
CC	in purified or natural biological materials. They may also be used for in				
CC	situ staining, fluorescence-activated cell sorting, Western blotting and				
CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can				
CC	be used for in vitro expansion of megakaryocytes and their committed				
CC	progenitors alone or in conjunction with additional cytokines				
XX					
SQ	Sequence 10 AA;				
Query Match	59.0%;	Score 46;	DB 4;	Length 10;	
Best Local Similarity	77.8%;	Pred. No. 3.8;			
Matches	7;	Conservative	1;	Mismatches	1; Indels 0; Gaps 0;
QY	1	CTLRWLHG	9		
	:				
Db	1	CTLRQWLQG	9		

RESULT 45
ABB72876
ID ABB72876 standard; peptide; 10 AA.
XX AC ABB72876;
XX DT
XX DT 05-APR-2002 (first entry)
DE TPO mimetic peptide SEQ ID NO:46.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
XX (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 5; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTLREWLHG 9
Db 1 CTLRQWLQG 9
RESULT 46
ADJ73027
ID ADJ73027 standard; peptide; 10 AA.
XX
AC ADJ73027;
XX
DT 06-MAY-2004 (first entry)
XX TPO mimetic peptide sequence SeqID 481.
DE
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX Synthetic.
OS
XX WO2003084477-A2.
PN
XX 16-OCT-2003.
PD
XX 24-MAR-2003; 2003WO-US009139.
PF
XX 29-MAR-2002; 2002US-0368791P.
PR
XX (CENZ) CENTOCOR INC.
PA
XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
PI WPI; 2003-804237/75.
DR
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 481; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;
Query Match 59.0%; Score 46; DB 7; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTLREWLHG 9
Db 1 CTLRQWLQG 9

```
RESULT 47
ADJ52662
ID ADJ52662 standard; peptide; 10 AA.
XX
AC
XX
AC ADJ52662;
DT 06-MAY-2004 (first entry)
XX
DE
XX
DE CH1 deleted mimetibody-related peptide SeqID481.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX WO2004002417-A2.
XX
XX 08-JAN-2004.
PD
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
PR
XX
XX (CENZ ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 481; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 8; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
| | | | |
Db 1 CTLRQLQG 9

RESULT 48
```

```
ADJ51623
ID ADJ51623 standard; peptide; 10 AA.
XX
AC ADJ51623;
XX
DT 06-MAY-2004 (first entry)
XX
DE
XX
DE CH1 deleted mimetibody-related peptide SeqID481.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
PD
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 481; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 8; Length 10;
```


Best Local Similarity 77.8%; Pred. No. 3.8;		Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 CTLREWLHG 9		
Db	1 CTLRQWLQG 9		
RESULT 49			
AAW09523			
ID	AAW09523 standard; protein; 12 AA.		
XX			
AC	AAW09523;		
XX			
DT	10-SEP-1997 (first entry)		
XX			
DE	Thrombopoietin receptor binding peptide.		
XX			
KW	Haematology; thrombocytopenia; TPO; TR; proliferation;		
KW	bone marrow transfusion; chemotherapy; radiation therapy.		
XX			
OS	Synthetic.		
XX			
PN	WO9640189-A1.		
XX			
PD	19-DEC-1996.		
XX			
PF	05-JUN-1996; 96WO-US008998.		
XX			
PR	07-JUN-1995; 95US-00472371.		
PR	07-JUN-1995; 95US-00473604.		
PR	07-JUN-1995; 95US-00476168.		
PR	07-JUN-1995; 95US-00478128.		
PR	07-JUN-1995; 95US-00484090.		
PR	07-JUN-1995; 95US-00485301.		
XX			
PA	(GLAX) GLAXO GROUP LTD.		
XX			
PI	Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;		
PI	Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;		
XX			
DR	WPI; 1997-051883/05.		
XX			
PT	Thrombopoietin receptor-binding/activating peptide(s) and peptide		
PT	mimetic(s) - useful in treatment of haematological disorders, esp.		
PT	thrombocytopenia resulting from chemotherapy, etc.		
XX			
PS	Disclosure; Page 27; 106pp; English.		
XX			
CC	The present sequence is a peptide which binds to thrombopoietin (TPO)		
CC	receptor (TR). The compound can be used for treating patients suffering		
CC	from haematological disorders and thrombocytopenia resulting from		
CC	chemotherapy, radiation therapy or bone marrow transfusions. The peptide		
CC	may also be used to maintain the proliferation and growth of TPO-		
CC	dependent cell lines and for use in biological research, for detecting		
CC	TPO receptors on living cells		
XX			
SQ	Sequence 12 AA;		
Query Match 59.0%; Score 46; DB 2; Length 12;		Best Local Similarity 58.3%; Pred. No. 4.5;	
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
Qy	1 CTLREWLHGGFC 12		
Db	1 CTLRQWILLGMC 12		
Search completed: May 12, 2006, 10:35:24			
Job time : 118.279 secs			
RESULT 50			
AAW36674			
ID	AAW36674 standard; peptide; 12 AA.		
XX			
AC	AAW36674;		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 14.5641 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTLREWLHGGFC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	47	60.3	97	1	OGBO6
2	45	57.7	361	2	S57895
3	45	57.7	503	2	B84228
4	45	57.7	532	1	CPBYI
5	44	56.4	171	2	G84512
6	43	55.1	791	2	A53691
7	42	53.8	380	2	T11335
8	41	52.6	218	2	D72073
9	41	52.6	218	2	D86550
10	41	52.6	309	2	E96566
11	41	52.6	530	2	A81958
12	41	52.6	531	2	E81015
13	41	52.6	619	2	F82984
14	41	52.6	1025	2	H86250
15	41	52.6	2225	1	A23443
16	41	52.6	4342	2	H83343
17	40.5	51.9	4391	2	A38096
18	40	51.3	117	2	B87354
19	40	51.3	139	2	S25970
20	40	51.3	312	2	G84305
21	40	51.3	378	2	S39007
22	40	51.3	379	1	S17406
23	40	51.3	379	1	S17408
24	40	51.3	379	1	S17410
25	40	51.3	379	1	S17415
26	40	51.3	379	1	S17417
27	40	51.3	379	1	S33572
28	40	51.3	379	1	S41832
29	40	51.3	379	1	S43261

30	40	51.3	379	1	S43262	ubiquinol-cytochro
31	40	51.3	379	1	S43263	ubiquinol-cytochro
32	40	51.3	379	1	S43264	ubiquinol-cytochro
33	40	51.3	379	1	S43265	ubiquinol-cytochro
34	40	51.3	379	1	S43266	ubiquinol-cytochro
35	40	51.3	379	1	S43267	ubiquinol-cytochro
36	40	51.3	379	1	S43269	ubiquinol-cytochro
37	40	51.3	379	2	E58889	ubiquinol-cytochro
38	40	51.3	379	2	I48133	ubiquinol-cytochro
39	40	51.3	379	2	I48132	ubiquinol-cytochro
40	40	51.3	379	2	I48180	ubiquinol-cytochro
41	40	51.3	379	2	E58851	ubiquinol-cytochro
42	40	51.3	379	2	T11375	ubiquinol-cytochro
43	40	51.3	379	2	T11869	ubiquinol-cytochro
44	40	51.3	379	2	T11492	ubiquinol-cytochro
45	40	51.3	379	2	S58990	ubiquinol-cytochro
46	40	51.3	379	2	T11401	ubiquinol-cytochro
47	40	51.3	379	2	A53077	ubiquinol-cytochro
48	40	51.3	379	2	I48134	ubiquinol-cytochro
49	40	51.3	380	1	CBRT	ubiquinol-cytochro
50	40	51.3	381	1	CBMS	ubiquinol-cytochro
51	40	51.3	381	2	T11440	ubiquinol-cytochro
52	40	51.3	382	1	S33573	ubiquinol-cytochro
53	40	51.3	382	2	S47882	ubiquinol-cytochro
54	40	51.3	438	2	C86244	DnaJ homolog, 4706
55	40	51.3	2180	2	T29764	hypothetical prote
56	40	51.3	2318	2	S45306	notch 3 protein -
57	40	51.3	2321	2	S78549	notch3 protein - h
58	40	51.3	4568	2	T08030	dynein beta heavy
59	39	50.0	151	2	D81814	hypothetical prote
60	39	50.0	160	2	H83047	hypothetical prote
61	39	50.0	209	2	AI3455	transcription regu
62	39	50.0	217	2	I48902	homeobox protein P
63	39	50.0	252	2	T29527	hypothetical prote
64	39	50.0	269	2	F71519	probable rRNA meth
65	39	50.0	372	2	T11100	ubiquinol-cytochro
66	39	50.0	379	1	S43268	ubiquinol-cytochro
67	39	50.0	379	1	S43270	ubiquinol-cytochro
68	39	50.0	384	2	S46110	hypothetical prote
69	39	50.0	492	2	T01086	probable serine/th
70	39	50.0	545	2	B44054	orf2 protein - Jun
71	39	50.0	726	2	T44187	infected cell prot
72	39	50.0	726	2	T44000	transport protein
73	39	50.0	861	2	A48825	Notch homolog Motc
74	39	50.0	899	2	G02428	subtilisin-like pr
75	39	50.0	915	1	A48225	subtilisin-like pr
76	39	50.0	915	2	JC6148	subtilisin-like pr
77	39	50.0	915	2	B48225	probable proprotei
78	39	50.0	1472	2	A26122	alpha-2-macroglobu
79	39	50.0	1474	1	MAHU	alpha-2-macroglobu
80	39	50.0	1476	2	JC5143	alpha-macroglobuli
81	39	50.0	1482	2	S13495	pregnancy zone pro
82	39	50.0	2531	2	S18188	notch protein homo
83	39	50.0	2531	2	A46019	notch-1 protein -
84	38.5	49.4	689	2	C84192	methionine-tRNA sy
85	38	48.7	112	2	D95269	hypothetical prote
86	38	48.7	141	2	D32536	T-cell receptor al
87	38	48.7	204	2	C83748	hypothetical prote
88	38	48.7	227	2	T49725	hypothetical prote
89	38	48.7	285	2	T22722	hypothetical prote
90	38	48.7	312	2	A87449	conserved hypothet
91	38	48.7	338	2	JC7161	acyl-CoA hydrolase
92	38	48.7	338	2	JC5415	palmitoyl-CoA hydr
93	38	48.7	343	2	JC5416	palmitoyl-CoA hydr
94	38	48.7	442	2	H86144	hypothetical prote
95	38	48.7	445	2	E84714	probable protein k
96	38	48.7	482	2	G86227	hypothetical prote
97	38	48.7	497	2	T15812	hypothetical prote
98	38	48.7	506	2	H75522	probable carboxyle
99	38	48.7	555	2	S63177	mannosyl transfera
100	38	48.7	631	2	T05103	hypothetical prote
101	38	48.7	658	1	A65079	arginine decarboxy
102	38	48.7	658	2	AG0876	arginine decarboxy

103 38 48.7 658 2 F91105 biosynthetic argin
104 38 48.7 658 2 A85951 biosynthetic argin
105 38 48.7 659 2 AB0114 arginine decarboxy
106 38 48.7 695 2 E87397 transglycosylase,
107 38 48.7 717 2 T28829 hypothetical prote
108 38 48.7 749 2 T31536 hypothetical prote
109 38 48.7 1057 2 T25396 hypothetical prote
110 38 48.7 1444 2 B84809 hypothetical prote
111 37.5 48.1 1274 2 E81779 proline dehydrogen
112 37 47.4 85 2 S01157 cytochrome-c oxida
113 37 47.4 97 2 I49360 cytochrome-c oxida
114 37 47.4 98 2 AC2369 hypothetical prote
115 37 47.4 140 2 E83910 hypothetical prote
116 37 47.4 167 2 I50487 ubiquinol-cytochro
117 37 47.4 170 2 B97811 proline/betaine tr
118 37 47.4 214 2 B81247 EpiH/GdmH-related
119 37 47.4 218 2 G82019 probable periplasm
120 37 47.4 230 2 AB3012 glutathione S-tran
121 37 47.4 230 2 E98272 glutathione transf
122 37 47.4 297 1 S49348 cytochrome-c oxida
123 37 47.4 302 2 JH0572 hypothetical prote
124 37 47.4 308 1 S22931 ubiquinol-cytochro
125 37 47.4 308 2 S22919 ubiquinol-cytochro
126 37 47.4 308 2 S22921 ubiquinol-cytochro
127 37 47.4 308 2 S22925 ubiquinol-cytochro
128 37 47.4 308 2 S22929 ubiquinol-cytochro
129 37 47.4 308 2 S22924 ubiquinol-cytochro
130 37 47.4 308 2 S22930 ubiquinol-cytochro
131 37 47.4 308 2 S22920 ubiquinol-cytochro
132 37 47.4 308 2 S22928 ubiquinol-cytochro
133 37 47.4 308 2 S22922 ubiquinol-cytochro
134 37 47.4 320 2 I51030 ubiquinol-cytochro
135 37 47.4 332 2 A29711 deacetoxycephalosp
136 37 47.4 339 2 C86978 hypothetical prote
137 37 47.4 339 2 F70901 riboflavin bifunct
138 37 47.4 344 2 A48990 transcription regu
139 37 47.4 379 2 D90617 cytochrome b limpo
140 37 47.4 379 2 D90621 cytochrome b limpo
141 37 47.4 379 2 T10998 ubiquinol-cytochro
142 37 47.4 379 2 D90625 cytochrome b limpo
143 37 47.4 379 2 D90615 ubiquinol-cytochro
144 37 47.4 379 2 T11178 ubiquinol-cytochro
145 37 47.4 379 2 T11530 ubiquinol-cytochro
146 37 47.4 379 2 D90613 cytochrome b limpo
147 37 47.4 379 2 D90627 cytochrome b limpo
148 37 47.4 379 2 D90619 cytochrome b limpo
149 37 47.4 380 1 D34285 ubiquinol-cytochro
150 37 47.4 380 1 S04840 ubiquinol-cytochro
151 37 47.4 380 1 S10198 ubiquinol-cytochro
152 37 47.4 380 2 T11033 ubiquinol-cytochro
153 37 47.4 380 2 T11803 ubiquinol-cytochro
154 37 47.4 380 2 I51374 ubiquinol-cytochro
155 37 47.4 380 2 E58893 ubiquinol-cytochro
156 37 47.4 380 2 T11466 ubiquinol-cytochro
157 37 47.4 380 2 T11204 ubiquinol-cytochro
158 37 47.4 380 2 T11191 ubiquinol-cytochro
159 37 47.4 380 2 S42245 ubiquinol-cytochro
160 37 47.4 380 2 T11086 ubiquinol-cytochro
161 37 47.4 381 2 S68140 ubiquinol-cytochro
162 37 47.4 384 2 S34816 nitrogenase cofact
163 37 47.4 420 2 T36193 probable salicylat
164 37 47.4 437 2 T18555 ATP-binding protei
165 37 47.4 477 2 A75052 cysteinyl-tRNA syn
166 37 47.4 591 2 G01586 probable protein 5
167 37 47.4 594 2 E96667 unknown protein, 6
168 37 47.4 644 2 A43370 capsid protein - N
169 37 47.4 701 2 S61239 hypothetical prote
170 37 47.4 754 2 S04381 beta-glucosidase (
171 37 47.4 850 2 S56015 gastric mucin MUC5
172 37 47.4 981 2 T50851 receptor protein k
173 37 47.4 984 2 E70406 DMSO reductase cha
174 37 47.4 987 2 T50850 receptor protein k
175 37 47.4 987 2 T50850

176 37 47.4 1023 2 E71376 conserved hypothet
177 37 47.4 1370 2 T19188 hypothetical prote
178 37 47.4 1378 2 G88637 protein F53H1.4 [i
179 37 47.4 3229 2 S27852 probable cell-surf
180 36.5 46.8 203 2 T24042 hypothetical prote
181 36.5 46.8 452 2 AG1223 cobyric acid a,c
182 36 46.2 100 2 S76498 ribosomal protein
183 36 46.2 100 2 AB2302 30S ribosomal prot
184 36 46.2 120 2 B90158 conserved hypothet
185 36 46.2 178 2 T29352 hypothetical prote
186 36 46.2 190 2 PQ0019 adenosinetriphosph
187 36 46.2 220 2 D87535 glutathione S-tran
188 36 46.2 225 2 B72749 probable proteasom
189 36 46.2 233 2 S52461 hypothetical prote
190 36 46.2 233 2 S38228 hypothetical prote
191 36 46.2 233 2 S53104 hypothetical prote
192 36 46.2 279 2 S75971 hypothetical prote
193 36 46.2 279 2 S62364 L-2-chloropropioni
194 36 46.2 301 2 S62416 n-acetylglucosamin
195 36 46.2 333 2 E97257 spore coat protein
196 36 46.2 345 2 T29786 hypothetical prote
197 36 46.2 345 2 T16869 hypothetical prote
198 36 46.2 370 2 JE0342 E2F transcription
199 36 46.2 377 2 C70628 hypothetical prote
200 36 46.2 379 2 I49399 ubiquinol-cytochro
201 36 46.2 427 2 S74211 PAS-6/7 protein pr
202 36 46.2 430 2 B83633 hypothetical prote
203 36 46.2 454 2 AI0725 para-aminobenzoate
204 36 46.2 454 2 A31132 p-aminobenzoate sy
205 36 46.2 470 2 E83947 hypothetical prote
206 36 46.2 471 2 D83546 probable amino aci
207 36 46.2 472 2 G83537 aromatic amino aci
208 36 46.2 473 2 B81007 cysteinyl-tRNA syn
209 36 46.2 520 2 T37957 probable cytochrom
210 36 46.2 617 2 S08317 hydrogenase (EC 1.
211 36 46.2 619 2 A45625 phosphoenolpyruvat
212 36 46.2 624 2 A83237 hypothetical prote
213 36 46.2 628 2 A55421 nucleoside-triphos
214 36 46.2 629 2 T37255 acetylcholinestera
215 36 46.2 646 2 S36586 E1 protein - human
216 36 46.2 666 2 T17396 vrlp protein - Dic
217 36 46.2 672 2 JC4637 transketolase (EC
218 36 46.2 672 2 T03457 transketolase (EC
219 36 46.2 699 2 H82030 cysteine-tRNA liga
220 36 46.2 788 2 T25061 hypothetical prote
221 36 46.2 788 2 T29768 hypothetical prote
222 36 46.2 860 2 C86203 hypothetical prote
223 36 46.2 973 2 AB2340 hypothetical prote
224 36 46.2 1119 2 T37460 probable sodium bi
225 36 46.2 1175 2 T22491 hypothetical prote
226 36 46.2 1201 2 F81202 proline dehydrogen
227 36 46.2 1230 2 T07663 soluble starch syn
228 36 46.2 1566 2 T20058 hypothetical prote
229 36 46.2 1584 2 T18276 protein-tyrosine k
230 36 46.2 1674 2 T01265 starch synthase DU
231 36 46.2 2531 2 T31070 notch homolog - se
232 35.5 45.5 292 2 H82955 hypothetical prote
233 35.5 45.5 298 2 G72531 hypothetical prote
234 35.5 45.5 394 2 A48100 B-type cyclin, Cig
235 35.5 45.5 411 2 S44344 cyclin cycl17 - fis
236 35.5 45.5 753 2 T19338 hypothetical prote
237 35.5 45.5 832 2 E84543 probable beta-gala
238 35 44.9 101 2 F72573 hypothetical prote
239 35 44.9 120 2 S10865 early E4 13K prote
240 35 44.9 145 2 S74292 hypothetical prote
241 35 44.9 169 2 T31484 hypothetical prote
242 35 44.9 174 2 G83712 hypothetical prote
243 35 44.9 185 2 F35719 hisp-like nucleoti
244 35 44.9 185 2 B86104 ATP-binding compon
245 35 44.9 185 2 E91263 ATP-binding compon
246 35 44.9 197 2 JC6544 tumor-associated a
247 35 44.9 206 2 D97285 ribosomal protein
248 35 44.9 216 2 T18176 hypothetical prote

249	35	44.9	228	2	JC4869	ribonuclease S4 (E	322	35	44.9	429	2	AD3642	(S)-2-hydroxy-acid
250	35	44.9	229	2	F70897	hypothetical prote	323	35	44.9	433	2	AE2658	glycolate oxidase
251	35	44.9	234	2	T46203	transcription fact	324	35	44.9	433	2	B97440	glycolate oxidase
252	35	44.9	242	2	G91139	probable peptidogl	325	35	44.9	442	2	T01731	hypothetical prote
253	35	44.9	242	2	B85985	probable peptidogl	326	35	44.9	449	2	C69079	probable phosphoma
254	35	44.9	242	2	B65112	Monofunctional bio	327	35	44.9	462	2	B81352	cysteine-tRNA liga
255	35	44.9	243	2	S38236	hypothetical prote	328	35	44.9	462	2	H69751	amino acid permeas
256	35	44.9	243	2	C95908	hypothetical prote	329	35	44.9	466	2	G83663	cysteiny1-tRNA syn
257	35	44.9	275	2	T10310	apoptosis-inhibiti	330	35	44.9	466	2	AG0839	Gaba permease (4-a
258	35	44.9	281	2	A43749	lignin beta-ether	331	35	44.9	466	2	A39599	55K erythrocyte me
259	35	44.9	299	2	B82535	conserved hypothet	332	35	44.9	481	2	T05270	probable amino aci
260	35	44.9	305	2	A46476	B cell-associated	333	35	44.9	494	2	D84860	probable serine/th
261	35	44.9	306	1	CPBOB	carboxypeptidase B	334	35	44.9	495	2	B96609	probable protein k
262	35	44.9	308	2	S22932	ubiquinol-cytochro	335	35	44.9	495	2	AE0666	probable esterase
263	35	44.9	308	2	S22923	ubiquinol-cytochro	336	35	44.9	502	2	T35910	probable carboxyle
264	35	44.9	312	2	H88567	protein K11H3.3 li	337	35	44.9	502	2	A39351	tetragonal surface
265	35	44.9	315	2	B98226	hypothetical prote	338	35	44.9	512	2	T47793	receptor-like prot
266	35	44.9	315	2	AF3060	conserved hypothet	339	35	44.9	512	2	VHN2P3	nucleocapsid prote
267	35	44.9	341	2	T07145	epoxide hydrolase	340	35	44.9	515	1	T18226	cytochrome P450 -
268	35	44.9	342	2	AD3450	proline racemase (341	35	44.9	517	2	T18226	nucleocapsid prote
269	35	44.9	351	2	I39808	spore coat protein	342	35	44.9	524	1	A48341	invasion protein i
270	35	44.9	361	2	T45918	hypothetical prote	343	35	44.9	563	2	S54420	hypothetical prote
271	35	44.9	366	2	T11271	ubiquinol-cytochro	344	35	44.9	604	2	G89864	probable peptidase
272	35	44.9	368	2	T26338	hypothetical prote	345	35	44.9	608	2	F83397	probable peptidase
273	35	44.9	370	2	H64480	hypothetical prote	346	35	44.9	618	2	S09251	hydrogenase (EC 1.
274	35	44.9	374	2	S40756	hypothetical prote	347	35	44.9	623	2	B81037	lipopolysaccharide
275	35	44.9	378	1	S17412	ubiquinol-cytochro	348	35	44.9	630	2	A81196	biosynthetic argin
276	35	44.9	378	2	D83381	hypothetical prote	349	35	44.9	630	2	D81831	arginine decarboxy
277	35	44.9	379	1	CBBO	ubiquinol-cytochro	350	35	44.9	673	2	T36717	probable serine/th
278	35	44.9	379	1	S17405	ubiquinol-cytochro	351	35	44.9	692	2	B64381	hypothetical prote
279	35	44.9	379	1	S17407	ubiquinol-cytochro	352	35	44.9	697	2	H85205	potassium channel
280	35	44.9	379	1	S17409	ubiquinol-cytochro	353	35	44.9	697	2	T04931	potassium channel
281	35	44.9	379	1	S17411	ubiquinol-cytochro	354	35	44.9	724	1	QQBEI5	UL89 protein - hum
282	35	44.9	379	1	S17413	ubiquinol-cytochro	355	35	44.9	801	2	A47744	diacylglycerol kin
283	35	44.9	379	1	S17414	ubiquinol-cytochro	356	35	44.9	823	2	S59141	mudrA protein - ma
284	35	44.9	379	1	S17418	ubiquinol-cytochro	357	35	44.9	952	2	E86147	T1N6.4 protein - A
285	35	44.9	379	1	S17419	ubiquinol-cytochro	358	35	44.9	1204	2	A96676	hypothetical prote
286	35	44.9	379	1	S17420	ubiquinol-cytochro	359	35	44.9	1267	1	MWXR32	lambda 3 protein -
287	35	44.9	379	2	I48135	ubiquinol-cytochro	360	35	44.9	1275	2	T38397	probable GTPase ac
288	35	44.9	379	2	I49400	ubiquinol-cytochro	361	35	44.9	1281	2	T15762	hypothetical prote
289	35	44.9	379	2	T11453	ubiquinol-cytochro	362	35	44.9	1328	2	T43060	agrin - electric r
290	35	44.9	379	2	T11505	ubiquinol-cytochro	363	35	44.9	1341	2	H98323	hypothetical prote
291	35	44.9	379	2	S58455	ubiquinol-cytochro	364	35	44.9	1426	2	E90456	oxydoreductase, pr
292	35	44.9	379	2	S58454	ubiquinol-cytochro	365	35	44.9	1444	2	T18856	angiogenesis inhib
293	35	44.9	379	2	T11414	ubiquinol-cytochro	366	35	44.9	1458	1	A49707	phospholipase A2 r
294	35	44.9	379	2	S58451	ubiquinol-cytochro	367	35	44.9	1597	1	BVFFSL	sol protein, large
295	35	44.9	379	2	S58452	ubiquinol-cytochro	368	35	44.9	1597	2	T08428	gene small optic l
296	35	44.9	379	2	S58457	ubiquinol-cytochro	369	35	44.9	1873	2	A55645	calcium channel, v
297	35	44.9	379	2	T11259	ubiquinol-cytochro	370	35	44.9	1959	1	AGRT	agrin - rat
298	35	44.9	379	2	T11349	ubiquinol-cytochro	371	35	44.9	3010	1	GNWVCJ	genome polypotein
299	35	44.9	379	2	S50334	ubiquinol-cytochro	372	35	44.9	3010	1	S18030	genome polypotein
300	35	44.9	379	2	S58462	ubiquinol-cytochro	373	35	44.9	3011	1	GNWVC3	genome polypotein
301	35	44.9	379	2	S58057	ubiquinol-cytochro	374	35	44.9	3011	1	GNWVCH	genome polypotein
302	35	44.9	379	2	S58085	ubiquinol-cytochro	375	35	44.9	3011	1	S40770	genome polypotein
303	35	44.9	379	2	S58456	ubiquinol-cytochro	376	35	44.9	5147	1	IJFFTM	cadherin-related t
304	35	44.9	379	2	S58459	ubiquinol-cytochro	377	35	44.9	6831	2	A88852	protein unc-22 lim
305	35	44.9	379	2	S58458	ubiquinol-cytochro	378	35	44.9	6839	2	S57242	twitchin [similar
306	35	44.9	379	2	S58460	ubiquinol-cytochro	379	35	44.9	7160	2	T27935	hypothetical prote
307	35	44.9	380	2	T11299	ubiquinol-cytochro	380	34.5	44.2	212	2	C87585	hypothetical prote
308	35	44.9	386	2	S72168	dopamine receptor	381	34.5	44.2	414	2	C69530	3-ketoacyl-CoA thi
309	35	44.9	391	2	E83151	hypothetical prote	382	34.5	44.2	1401	2	T48079	hypothetical prote
310	35	44.9	395	2	F71118	hypothetical prote	383	34	43.6	91	2	G83144	hypothetical prote
311	35	44.9	397	2	AE0148	conserved hypothet	384	34	43.6	97	1	OGHU6A	cytochrome-c oxida
312	35	44.9	397	2	B81839	probable periplasm	385	34	43.6	102	2	C81682	conserved hypothet
313	35	44.9	409	2	D75629	probable transposa	386	34	43.6	105	2	T35995	probable dioxygena
314	35	44.9	415	2	A32129	carboxypeptidase B	387	34	43.6	106	2	JQ0234	hypothetical 12.5K
315	35	44.9	416	1	A42332	carboxypeptidase B	388	34	43.6	114	2	T11095	NADH2 dehydrogenas
316	35	44.9	416	2	D75407	probable transposa	389	34	43.6	116	2	A71565	hypothetical prote
317	35	44.9	417	1	A34487	carboxypeptidase A	390	34	43.6	137	2	B96603	transcription fact
318	35	44.9	423	2	A41204	carboxypeptidase B	391	34	43.6	138	2	G69328	Brute force ORF -
319	35	44.9	423	2	D90927	hypothetical prote	392	34	43.6	143	2	H70601	hypothetical prote
320	35	44.9	423	2	A64926	ynnC protein - Esc	393	34	43.6	166	2	H64416	VPS29-like phospho
321	35	44.9	423	2	H85775	hypothetical prote	394	34	43.6	169	2	D72649	hypothetical prote

395	34	43.6	185	2	E70029	conserved hypothet	468	34	43.6	511	2	E64183	xylulokinase homol
396	34	43.6	186	2	F90190	conserved hypothet	469	34	43.6	511	2	A69369	glutamate synthase
397	34	43.6	211	2	C82381	pyridoxamine 5'-ph	470	34	43.6	513	2	C58888	cytochrome-c oxida
398	34	43.6	243	2	G98156	hypothetical prote	471	34	43.6	513	2	T11443	cytochrome-c oxida
399	34	43.6	248	1	JQ1682	infected cell prote	472	34	43.6	513	2	T11482	cytochrome-c oxida
400	34	43.6	259	2	AG2462	hypothetical prote	473	34	43.6	513	2	T11391	cytochrome-c oxida
401	34	43.6	261	2	G88469	protein C28H8.10 [474	34	43.6	514	1	ODMS1	cytochrome-c oxida
402	34	43.6	267	2	I72882	Fc gamma receptor	475	34	43.6	514	2	T11495	cytochrome-c oxida
403	34	43.6	267	2	I56110	Fc gamma RIIIB-alp	476	34	43.6	514	2	T11142	cytochrome-c oxida
404	34	43.6	267	2	A35902	Fc gamma (IGG) rec	477	34	43.6	514	2	T11404	cytochrome-c oxida
405	34	43.6	290	2	G65031	hypothetical prote	478	34	43.6	514	2	S41837	cytochrome-c oxida
406	34	43.6	290	2	E85899	hypothetical prote	479	34	43.6	514	2	S26153	cytochrome-c oxida
407	34	43.6	290	2	B91055	hypothetical prote	480	34	43.6	514	2	T11365	cytochrome-c oxida
408	34	43.6	292	2	T02349	hypothetical prote	481	34	43.6	514	2	T11859	cytochrome-c oxida
409	34	43.6	297	2	S55063	conserved hypothet	482	34	43.6	514	2	T11249	cytochrome-c oxida
410	34	43.6	298	2	AD2715	ABC transporter, m	483	34	43.6	514	2	T10974	cytochrome-c oxida
411	34	43.6	299	2	AH0245	probable 4-diphosp	484	34	43.6	514	2	S04749	cytochrome-c oxida
412	34	43.6	301	2	H70505	hypothetical prote	485	34	43.6	515	1	VHNZB3	nucleocapsid prote
413	34	43.6	306	2	A72668	probable spermidin	486	34	43.6	515	2	C58892	cytochrome-c oxida
414	34	43.6	308	2	S22926	ubiquinol-cytochro	487	34	43.6	515	2	T11339	cytochrome-c oxida
415	34	43.6	308	2	A84566	Ac-like transposas	488	34	43.6	516	2	S36008	cytochrome-c oxida
416	34	43.6	308	2	S58504	reverse transcript	489	34	43.6	516	2	S35464	cytochrome-c oxida
417	34	43.6	309	2	A38395	mast cell carboxyp	490	34	43.6	516	2	T09859	cytochrome-c oxida
418	34	43.6	313	2	T35935	transcription regu	491	34	43.6	516	2	T09949	cytochrome-c oxida
419	34	43.6	315	2	T18696	hypothetical prote	492	34	43.6	516	2	S45351	cytochrome-c oxida
420	34	43.6	320	2	AF1892	[NiFe] uptake hydr	493	34	43.6	517	2	S55006	cytochrome-c oxida
421	34	43.6	321	2	S58686	dTDPglucose 4,6-de	494	34	43.6	517	2	T11302	cytochrome-c oxida
422	34	43.6	322	1	HLHUR2	T-cell surface gly	495	34	43.6	518	2	T11536	cytochrome-c oxida
423	34	43.6	328	1	DWSMGG	dTDPglucose 4,6-de	496	34	43.6	518	2	T11456	cytochrome-c oxida
424	34	43.6	334	2	I59348	CCAAT binding tran	497	34	43.6	518	2	T11766	cytochrome-c oxida
425	34	43.6	335	2	T51106	dTDPglucose 4,6-de	498	34	43.6	518	2	T11289	cytochrome-c oxida
426	34	43.6	353	1	C69979	conserved hypothet	499	34	43.6	521	2	T11166	CPDdiacylglycerol-
427	34	43.6	358	1	WMBE38	infected cell prot	500	34	43.6	522	2	T45861	hypothetical prote
428	34	43.6	360	2	S54256	photosystem II pro	501	34	43.6	523	2	T22728	hypothetical prote
429	34	43.6	371	2	PS0016	5-methyltetrahydro	502	34	43.6	524	1	VHNZP1	nucleocapsid prote
430	34	43.6	375	1	YKMY	citrate (si)-synth	503	34	43.6	524	1	VHNZT1	nucleocapsid prote
431	34	43.6	378	1	C30020	ubiquinol-cytochro	504	34	43.6	544	2	B83347	hypothetical prote
432	34	43.6	378	1	S01190	ubiquinol-cytochro	505	34	43.6	547	2	D84942	methionine-tRNA li
433	34	43.6	379	2	S58466	ubiquinol-cytochro	506	34	43.6	568	1	RRNZ35	polymerase-associa
434	34	43.6	380	2	A71390	ubiquinol-cytochro	507	34	43.6	568	1	RRNZ39	polymerase-associa
435	34	43.6	388	1	CBZM	ubiquinol-cytochro	508	34	43.6	571	1	JN0858	chitinase (EC 3.2.
436	34	43.6	388	2	T29364	hypothetical prote	509	34	43.6	571	2	T42071	probable chitinase
437	34	43.6	392	1	CEVF	ubiquinol-cytochro	510	34	43.6	584	2	I50419	s-gicerin precurs
438	34	43.6	393	1	CBPOM	ubiquinol-cytochro	511	34	43.6	597	2	AH2351	serine/threonine k
439	34	43.6	393	1	S38960	ubiquinol-cytochro	512	34	43.6	603	2	H90061	hypothetical prote
440	34	43.6	394	1	CBOBE	ubiquinol-cytochro	513	34	43.6	609	2	AH0388	alpha-glucosidase
441	34	43.6	397	1	CBRZ	ubiquinol-cytochro	514	34	43.6	615	1	KFHU12	coagulation factor
442	34	43.6	397	2	T14263	ubiquinol-cytochro	515	34	43.6	616	2	T29234	hypothetical prote
443	34	43.6	398	1	A22931	hypothetical prote	516	34	43.6	621	2	T06717	hypothetical prote
444	34	43.6	403	2	H87569	ubiquinol-cytochro	517	34	43.6	621	2	B95897	probable cellulose
445	34	43.6	404	1	S25953	biphenyl dioxygena	518	34	43.6	647	1	QYFFGM	phosphoenolpyruvat
446	34	43.6	406	1	JC4996	probable protease	519	34	43.6	656	2	C711361	probable long-chai
447	34	43.6	406	2	A97266	collagenase family	520	34	43.6	661	2	T37753	hypothetical prote
448	34	43.6	406	2	B97108	glycolate oxidase	521	34	43.6	676	1	A39379	hatching-suppresse
449	34	43.6	408	2	C82977	ccaa-binding fact	522	34	43.6	727	2	A55628	translation initia
450	34	43.6	415	2	T40338	carboxypeptidase A	523	34	43.6	762	2	G88436	protein T04A8.13 [
451	34	43.6	417	2	A43929	hypothetical prote	524	34	43.6	791	2	T24435	hypothetical prote
452	34	43.6	420	2	B87201	molybdopterin bios	525	34	43.6	799	2	T02981	1,4-alpha-glucan b
453	34	43.6	423	2	E84027	probable acyl-CoA	526	34	43.6	799	2	T01663	1,4-alpha-glucan b
454	34	43.6	425	2	G83046	conserved hypothet	527	34	43.6	807	2	A38152	F-spondin - rat
455	34	43.6	438	2	AD0437	probable flavonol	528	34	43.6	809	2	E90016	hyaluronate lyase
456	34	43.6	444	2	G84565	hypothetical 49.3K	529	34	43.6	877	1	A25962	mitosis inhibitor
457	34	43.6	450	2	E98303	nitriolotriacetate	530	34	43.6	879	2	F85875	probable fimbrial
458	34	43.6	450	2	AH2979	ribulose-bisphosph	531	34	43.6	879	2	E91031	probable outer mem
459	34	43.6	473	1	RKBLT	probable amino-aci	532	34	43.6	907	2	T27317	hypothetical prote
460	34	43.6	475	2	G81227	hypothetical prote	533	34	43.6	1022	2	S50534	hypothetical prote
461	34	43.6	477	2	S77373	protein F3M18.4 [i	534	34	43.6	1045	2	S23570	pol polypeptin ho
462	34	43.6	485	2	AE2165	hypothetical prote	535	34	43.6	1138	2	H84529	hypothetical prote
463	34	43.6	486	2	B86411	integral membrane	536	34	43.6	1224	2	S25952	gene cob intron 3
464	34	43.6	494	2	G97496	D-serine/D-alanine	537	34	43.6	1227	1	XYECMH	5-methyltetrahydro
465	34	43.6	497	2	AD3514	cytochrome-c oxida	538	34	43.6	1227	2	AH1011	5-methyltetrahydro
466	34	43.6	508	2	S74200	probable glu-tRNA	539	34	43.6	1227	2	G86093	hypothetical prote
467	34	43.6	509	2	A71254		540	34	43.6	1227	2	A98246	hypothetical prote

541	34	43.6	1237	2	T45070	protein kinase hom	614	33	42.3	288	2	B70715	hypothetical prote
542	34	43.6	1300	2	I53799	CG1 protein - huma	615	33	42.3	294	2	T10858	probable carboxype
543	34	43.6	1314	2	H86327	protein F18O14.25	616	33	42.3	294	2	T38758	60s ribosomal prot
544	34	43.6	1356	2	S32763	kinectin 1 - human	617	33	42.3	294	2	T39325	60s ribosomal prot
545	34	43.6	1416	2	E88550	protein ZC84.1 [im	618	33	42.3	297	1	JN0773	calponin H1 - rat
546	34	43.6	1637	2	T00070	hypothetical prote	619	33	42.3	297	1	S31484	calponin H1 - pig
547	34	43.6	1734	2	A54602	microtubule-associ	620	33	42.3	297	1	S31486	calponin H1 - mous
548	34	43.6	1955	1	AGCH	agrin precursor -	621	33	42.3	297	2	JC4500	basic calponin - h
549	34	43.6	1974	2	T30010	hypothetical prote	622	33	42.3	297	2	G02142	smooth muscle cell
550	34	43.6	2233	1	ZLN2P3	genome polypeptin	623	33	42.3	301	2	C82394	probable glycine c
551	34	43.6	2844	2	S28291	hypothetical prote	624	33	42.3	308	1	H65050	probable ATP-bindi
552	34	43.6	6260	2	T30228	polyketide synthas	625	33	42.3	308	1	C85919	hypothetical prote
553	34	43.6	8563	2	T30226	polyketide synthas	626	33	42.3	308	2	D91074	probable ATP-bindi
554	33.5	42.9	296	2	T06025	transcription fact	627	33	42.3	308	2	S74834	hypothetical prote
555	33.5	42.9	310	2	T46035	transcription fact	628	33	42.3	308	2	JC7125	epidermal growth f
556	33.5	42.9	321	1	E64026	probable S-adenosy	629	33	42.3	310	2	G86490	probable biotin ho
557	33.5	42.9	333	2	T51650	probable transcript	630	33	42.3	310	2	T33205	hypothetical prote
558	33.5	42.9	348	2	E96620	protein T30E16.21	631	33	42.3	311	2	T43794	probable lipase li
559	33.5	42.9	374	2	T05891	myb-related protei	632	33	42.3	311	2	F90116	guanine nucleotide
560	33.5	42.9	376	2	T51673	myb-related transc	633	33	42.3	321	2	AB0845	probable phosphosu
561	33.5	42.9	379	2	T16213	APX-1 protein homo	634	33	42.3	325	2	F96588	hypothetical prote
562	33.5	42.9	701	2	T16607	hypothetical prote	635	33	42.3	326	2	B95941	probable NADPH2:qu
563	33.5	42.9	1036	2	T05687	beta-galactosidase	636	33	42.3	328	2	T02046	prolamin box bindi
564	33.5	42.9	1151	2	H71347	hypothetical prote	637	33	42.3	329	2	T09897	hypothetical prote
565	33.5	42.9	1523	2	T13953	MEGF5 protein - ra	638	33	42.3	329	2	T14760	hypothetical prote
566	33.5	42.9	3430	1	GNWVWV	genome polypeptin	639	33	42.3	331	2	T20593	hypothetical prote
567	33.5	42.9	3432	1	GNWVJS	genome polypeptin	640	33	42.3	333	2	S49054	hypothetical prote
568	33.5	42.9	3432	1	GNWVJE	genome polypeptin	641	33	42.3	333	2	B42476	hypothetical prote
569	33.5	42.9	3433	1	GNWVKV	genome polypeptin	642	33	42.3	333	2	AC0766	hypothetical prote
570	33	42.3	97	1	D44001	vpr protein - huma	643	33	42.3	337	2	E69136	probable glycosylt
571	33	42.3	97	2	S51924	hypothetical prote	644	33	42.3	339	2	H71361	hypothetical prote
572	33	42.3	100	2	JQ0859	hypothetical 11K p	645	33	42.3	340	2	T44739	conserved hypothet
573	33	42.3	101	2	S26314	Ig heavy chain V r	646	33	42.3	342	2	AD3197	citrate synthase I
574	33	42.3	109	2	AF0682	probable membrane	647	33	42.3	344	2	T31913	hypothetical prote
575	33	42.3	126	2	H70634	dnpD protein - Str	648	33	42.3	345	2	S72833	periplasmic bindin
576	33	42.3	153	2	D24372	restriction system	649	33	42.3	347	2	H89835	hypothetical prote
577	33	42.3	153	2	F98079	DpnD protein limpo	650	33	42.3	349	2	H85626	hypothetical prote
578	33	42.3	153	2	H95215	molybdopterin bios	651	33	42.3	349	2	E90819	hypothetical prote
579	33	42.3	157	2	B69659	hypothetical prote	652	33	42.3	349	2	E90912	hypothetical prote
580	33	42.3	163	2	G90437	cytochrome-c oxida	653	33	42.3	367	2	S59146	ubiquinol-cytochro
581	33	42.3	184	2	F69273	hypothetical prote	654	33	42.3	367	2	T52363	hypothetical prote
582	33	42.3	193	1	JN0728	DNA binding protei	655	33	42.3	371	2	E72665	hypothetical prote
583	33	42.3	195	2	T46072	hypothetical prote	656	33	42.3	372	2	A70302	rod shape determin
584	33	42.3	204	1	QXBP9L	unknown protein en	657	33	42.3	372	2	AB3466	mandelate racemase
585	33	42.3	207	2	D90730	hypothetical prote	658	33	42.3	377	2	T11361	ubiquinol-cytochro
586	33	42.3	210	2	T10647	probable phosphat	659	33	42.3	378	2	B59180	Wnt inhibitory fac
587	33	42.3	213	2	E70810	phosphate phoyl pro	660	33	42.3	380	2	G97828	hypothetical prote
588	33	42.3	221	2	G70533	phosphate transpor	661	33	42.3	381	2	S60623	ubiquinol-cytochro
589	33	42.3	222	2	G87182	hypothetical prote	662	33	42.3	382	2	T11138	ubiquinol-cytochro
590	33	42.3	224	2	T32606	hypothetical prote	663	33	42.3	383	2	S49928	F420-nonreducing h
591	33	42.3	226	2	G86641	hypothetical prote	664	33	42.3	383	2	S67477	F420-nonreducing h
592	33	42.3	228	2	T01350	phosphoglycerate m	665	33	42.3	385	1	S01511	ubiquinol-cytochro
593	33	42.3	230	2	C95028	phosphoglycerate m	666	33	42.3	399	2	C83611	hypothetical prote
594	33	42.3	231	2	E86743	hypothetical prote	667	33	42.3	403	2	B69531	dihydroorotase (py
595	33	42.3	232	2	T09898	hypothetical prote	668	33	42.3	409	2	T32607	hypothetical prote
596	33	42.3	232	2	T01867	hypothetical prote	669	33	42.3	414	2	H97601	aminotransferase I
597	33	42.3	233	2	A84777	hypothetical prote	670	33	42.3	414	2	AI2823	hypothetical prote
598	33	42.3	239	2	AE2410	hypothetical prote	671	33	42.3	417	2	S07286	aminotransferase I
599	33	42.3	244	2	T31838	ABC-type transport	672	33	42.3	418	2	G83904	hypothetical prote
600	33	42.3	247	2	T00107	ribosomal protein	673	33	42.3	428	2	B82248	long-chain fatty a
601	33	42.3	248	2	T43382	phosphoglycerate m	674	33	42.3	432	2	A85040	hypothetical prote
602	33	42.3	248	2	C97899	phosphoglycerate m	675	33	42.3	432	2	T01925	hypothetical prote
603	33	42.3	249	2	T04939	hypothetical prote	676	33	42.3	443	2	F82231	GGDEF family prote
604	33	42.3	249	2	T21920	hypothetical prote	677	33	42.3	446	2	T35094	hypothetical prote
605	33	42.3	251	2	AB2352	probable phosphopr	678	33	42.3	447	2	G95068	cysteinyI-trRNA syn
606	33	42.3	259	2	C71430	hypothetical prote	679	33	42.3	448	2	G97936	cysteine-tRNA liga
607	33	42.3	262	2	E84694	probable tropinone	680	33	42.3	448	2	C86856	probable serine/th
608	33	42.3	262	2	E84694	hypothetical prote	681	33	42.3	450	2	T01711	p-aminobenzoate sy
609	33	42.3	263	2	T33290	hypothetical prote	682	33	42.3	453	1	AGEC1	p-aminobenzoate sy
610	33	42.3	264	2	F89833	teichoic acid tran	683	33	42.3	453	2	E85792	p-aminobenzoate sy
611	33	42.3	272	2	H95314	probable transposa	684	33	42.3	454	2	A99944	similar to phospho
612	33	42.3	278	2	F84127	hypothetical prote	685	33	42.3	457	2	D71717	cysteine-tRNA liga
613	33	42.3					686	33	42.3				

687 33 42.3 461 2 A97491 cysteinyl-tRNA syn
688 33 42.3 461 2 AH2708 cysteinyl-tRNA syn
689 33 42.3 463 2 C87306 cysteinyl-tRNA syn
690 33 42.3 463 2 B56849 dopamine receptor-
691 33 42.3 465 1 F64630 cysteine-tRNA liga
692 33 42.3 465 2 T16835 hypothetical prote
693 33 42.3 465 2 D71884 cysteine-tRNA liga
694 33 42.3 466 2 C89820 cysteinyl-tRNA syn
695 33 42.3 469 2 B70607 probable cyss prot
696 33 42.3 470 2 G86516 dicarboxylase tran
697 33 42.3 470 2 H72105 dicarboxylase tran
698 33 42.3 471 2 G83561 hypothetical prote
699 33 42.3 472 1 RKRL1 ribulose-bisphosph
700 33 42.3 473 2 S18315 ribulose-bisphosph
701 33 42.3 473 2 C86949 probable cysteinyl
702 33 42.3 476 2 D71108 cysteine-tRNA liga
703 33 42.3 476 2 G84634 probable prolylcar
704 33 42.3 489 2 B69680 para-nitrobenzyl e
705 33 42.3 491 2 S74473 probable starch sy
706 33 42.3 493 2 A45737 phenylcarbamate hy
707 33 42.3 497 2 A86146 hypothetical prote
708 33 42.3 501 2 T27613 hypothetical prote
709 33 42.3 503 2 C71535 probable NADH (ubi
710 33 42.3 506 2 AB3411 cysteine-tRNA liga
711 33 42.3 507 1 C53308 probable hydro-lya
712 33 42.3 509 2 A91024 NADH dehydrogenase
713 33 42.3 509 2 C64999 NADH2 dehydrogenase
714 33 42.3 509 2 B85868 NADH dehydrogenase
715 33 42.3 520 2 F69470 dipeptide ABC tran
716 33 42.3 523 2 JC7556 linoleoyl-CoA desa
717 33 42.3 532 2 A34329 60K esterase (EC 3
718 33 42.3 534 2 H82244 sensor histidine k
719 33 42.3 539 2 A29923 carboxylesterase (
720 33 42.3 540 2 A75250 carboxylesterase,
721 33 42.3 540 2 A31584 carboxylesterase (
722 33 42.3 543 2 S35047 mucin JUL7 - human
723 33 42.3 547 2 AE3017 conserved hypothet
724 33 42.3 547 2 C98267 hypothetical prote
725 33 42.3 549 2 JX0054 carboxylesterase (
726 33 42.3 554 1 S34607 carboxylesterase (
727 33 42.3 554 2 A39060 carboxylesterase (
728 33 42.3 554 2 G75576 conserved hypothet
729 33 42.3 557 2 A47162 thiolesterase B (E
730 33 42.3 559 1 JC5408 carboxylesterase (
731 33 42.3 561 2 S71597 carboxylesterase (
732 33 42.3 561 2 S47655 carboxylesterase (
733 33 42.3 561 2 JC2447 carboxylesterase (
734 33 42.3 561 2 S62788 carboxylesterase (
735 33 42.3 562 2 S27800 elastase precursor
736 33 42.3 562 2 A55281 carboxylesterase (
737 33 42.3 563 2 D84717 probable cysteinyl
738 33 42.3 565 2 S10367 carboxylesterase (
739 33 42.3 566 2 S19307 carboxylesterase (
740 33 42.3 567 1 A41010 carboxylesterase (
741 33 42.3 570 2 T32061 gut-specific carbo
742 33 42.3 571 2 T10232 hypothetical prote
743 33 42.3 571 2 AI0506 probable sulfatase
744 33 42.3 580 2 T18439 hypothetical prote
745 33 42.3 583 2 C69158 sensory transducti
746 33 42.3 588 2 E87521 peptidase M1 famil
747 33 42.3 596 1 HQ2JUL hydrogenase (EC 1.
748 33 42.3 596 1 S11969 hydrogenase (EC 1.
749 33 42.3 600 2 T18446 hypothetical prote
750 33 42.3 601 1 S11777 hydrogenase (EC 1.
751 33 42.3 602 1 JQ0806 hydrogenase (EC 1.
752 33 42.3 606 2 A70960 probable pckA prot
753 33 42.3 614 2 G71298 probable methyl-ac
754 33 42.3 617 2 T29660 hypothetical prote
755 33 42.3 622 2 G81981 probable lipopolys
756 33 42.3 644 2 C70585 hypothetical prote
757 33 42.3 667 2 T23010 hypothetical prote
758 33 42.3 670 2 AH2425 hypothetical prote
759 33 42.3 681 2 S76354 ABC1-type transpor

760 33 42.3 682 1 JH0560 cyclic nucleotide-
761 33 42.3 707 2 T01502 probable serine/th
762 33 42.3 713 2 AI1752 endopeptidase [bac
763 33 42.3 716 1 BVECAD membrane protein t
764 33 42.3 738 2 S10659 membrane protein t
765 33 42.3 748 2 B45046 basic juvenile hor
766 33 42.3 757 2 B75437 ABC transporter, A
767 33 42.3 761 2 H65083 glycolate oxidase
768 33 42.3 785 2 A29953 alpha-1 proteinase
769 33 42.3 788 1 I59282 diacylglycerol kin
770 33 42.3 811 2 PN0689 connectin 1 - chic
771 33 42.3 828 2 S52393 beta-galactosidase
772 33 42.3 839 2 C84685 probable beta-gala
773 33 42.3 840 2 G72468 hypothetical prote
774 33 42.3 849 2 G84533 hypothetical prote
775 33 42.3 858 2 JC2309 chitin synthase (E
776 33 42.3 858 2 JC2308 chitin synthase (E
777 33 42.3 859 2 S62441 chitin synthase (E
778 33 42.3 877 2 S49197 envelope protein p
779 33 42.3 889 2 JC6015 chitin synthase (E
780 33 42.3 895 2 B64238 isoleucine-tRNA li
781 33 42.3 904 2 A84212 hypothetical prote
782 33 42.3 908 2 T22376 hypothetical prote
783 33 42.3 910 2 JC4609 chitin synthase (E
784 33 42.3 911 2 JC6016 chitin synthase (E
785 33 42.3 916 2 JC2315 chitin synthase (E
786 33 42.3 923 2 C97487 hypothetical prote
787 33 42.3 923 2 AC2705 conserved hypothet
788 33 42.3 944 2 T47246 chitin synthase (E
789 33 42.3 955 2 E84022 hypothetical prote
790 33 42.3 956 2 JH0826 glutamate ionotrop
791 33 42.3 956 2 JS0685 glutamate receptor
792 33 42.3 960 2 A41638 chitin synthase (E
793 33 42.3 979 2 JH0589 glutamate receptor
794 33 42.3 979 2 JH0592 glutamate receptor
795 33 42.3 980 2 I57936 glutamate receptor
796 33 42.3 980 2 T27342 hypothetical prote
797 33 42.3 983 2 A59054 chitin synthase (E
798 33 42.3 1005 2 A42265 alpha-mannosidase
799 33 42.3 1009 2 S20538 chitin synthase (E
800 33 42.3 1013 2 JC2314 chitin synthase (E
801 33 42.3 1013 2 T10659 probable serine/th
802 33 42.3 1036 2 B83466 probable RND efflu
803 33 42.3 1049 1 S51784 toxin III - Actino
804 33 42.3 1131 2 A23944 chitin synthase (E
805 33 42.3 1217 2 C86159 hypothetical prote
806 33 42.3 1231 2 C84716 hypothetical prote
807 33 42.3 1257 2 S28764 neurocan precursor
808 33 42.3 1268 2 S52781 neurocan - mouse
809 33 42.3 1295 2 S60179 pol polyprotein ho
810 33 42.3 1401 2 T39225 MAP kinase kinase
811 33 42.3 1421 2 T00333 hypothetical prote
812 33 42.3 1451 2 B41185 alpha-2 macroglobu
813 33 42.3 1464 2 JC5144 murinoglobulin pre
814 33 42.3 1473 2 A20872 ovostatin precurs
815 33 42.3 1476 2 A41185 alpha-2 macroglobu
816 33 42.3 1477 2 A29952 alpha-1 proteinase
817 33 42.3 1487 2 S15904 alpha-1 proteinase
818 33 42.3 1495 2 S27001 alpha-2-macroglobu
819 33 42.3 1500 2 A42210 alpha-1-macroglobu
820 33 42.3 1574 2 T13954 MEGF6 protein - ra
821 33 42.3 1743 2 T26859 hypothetical prote
822 33 42.3 1873 2 A30063 dihydropyridine re
823 33 42.3 1963 1 MWKW myosin heavy chain
824 33 42.3 1992 1 S02771 myosin heavy chain
825 33 42.3 2870 2 A35548 319K protein ndvB
826 33 42.3 3071 2 T50345 vacuolar protein s
827 33 42.3 4162 2 T42633 connectin/titin -
828 33 42.3 26926 1 I38344 titin, cardiac mus
829 32.5 41.7 80 2 H82194 conserved hypothet
830 32.5 41.7 154 2 T17816 hypothetical prote
831 32.5 41.7 191 2 AC3413 holo-lacyl-carrier
832 32.5 41.7 224 2 T21788 hypothetical prote

979	32	41.0	380	2	Tl11113	ubiquinol-cytochro
980	32	41.0	380	2	T09869	ubiquinol-cytochro
981	32	41.0	380	2	T09959	ubiquinol-cytochro
982	32	41.0	380	2	D90623	cytochrome b [limpo
983	32	41.0	381	2	Tl1546	ubiquinol-cytochro
984	32	41.0	381	2	Tl1312	ubiquinol-cytochro
985	32	41.0	381	2	Tl1776	ubiquinol-cytochro
986	32	41.0	381	2	T09640	protein phosphatas
987	32	41.0	383	2	Tl5043	fungal elicitor-in
988	32	41.0	384	2	JC5206	hypothetical 43.0K
989	32	41.0	386	2	Tl1832	ubiquinol-cytochro
990	32	41.0	387	2	F82815	voltage-gated pota
991	32	41.0	388	2	AB2372	hypothetical prote
992	32	41.0	389	2	A48966	methylase - Lactoc
993	32	41.0	397	2	D70512	hypothetical prote
994	32	41.0	398	1	S35473	ubiquinol-cytochro
995	32	41.0	399	2	G69868	hypothetical prote
996	32	41.0	400	2	AG1879	hypothetical prote
997	32	41.0	409	2	F90825	probable integrase
998	32	41.0	413	2	B84109	ABC transporter (A
999	32	41.0	415	2	T32467	hypothetical prote
1000	32	41.0	416	2	A85684	probable integrase

ALIGNMENTS

RESULT 1
OGB06
cytochrome-c oxidase (EC 1.9.3.1) chain VIA precursor, cardiac [validated] - bovine
N;Alternate names: cytochrome-c oxidase chain VIB
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Dec-1993 #sequence_revision 07-Jun-1996 #text_change 31-Dec-2004
C;Accession: S35702; S16242; A24659
R;Smith, E.O.; Lomax, M.I.
Biochim. Biophys. Acta 1174, 63-71, 1993
A;Title: Structural organization of the bovine gene for the heart/muscle isoform of cytochrome c oxidase
A;Reference number: S35702; MUID:93326637; PMID:7687470
A;Accession: S35702
A;Molecule type: DNA
A;Residues: 1-97 <SM2>
A;Cross-references: UNIPROT:P07471; UNIPARC:UPI000012813D; GB:S64127; NID:G404386; PIDN:R;Smith, E.O.; BeMent, D.M.; Grossman, L.I.; Lomax, M.I.
Biochim. Biophys. Acta 1089, 266-268, 1991
A;Title: The cDNA for the heart/ muscle isoform of bovine cytochrome c oxidase subunit V
A;Reference number: S16242; MUID:91274363; PMID:1647214
A;Accession: S16242
A;Molecule type: mRNA
A;Residues: 1-97 <SM1>
A;Cross-references: UNIPARC:UPI000012813D; EMBL:X56857; NID:G269; PIDN:CAA40183.1; PID:SR;Meinecke, L.; Buse, G.
Biol. Chem. Hoppe-Seyler 366, 687-694, 1985
A;Title: Studies on cytochrome c oxidase. XII. Isolation and primary structure of polypeptide chain of bovine heart cytochrome c oxidase
A;Reference number: A24659; MUID:86000134; PMID:2994692
A;Accession: A24659
A;Molecule type: protein
A;Residues: 13-96 <MEI>
A;Cross-references: UNIPARC:UPI000011289B
A;Experimental source: heart
R;Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-ito, submitted to the Brookhaven Protein Data Bank, April 1996
A;Reference number: A67451; PDB:1OCC
A;Contents: annotation; X-ray crystallography, 2.8 angstroms, residues 13-96
R;Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K Science 272, 1136-1144, 1996
A;Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 angstroms
A;Reference number: A57981; MUID:96216288; PMID:8638158
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
C;Comment: This form is expressed in fetal and adult cardiac and skeletal muscle.
C;Genetics:
A;Gene: COX6A1
A;Introns: 25/1; 70/3

C;Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and consists of two subunits, Vb (see PIR:OGB06A), V1a, V1b (see PIR:OGB07), V1c (see PIR:OGB06C), V11a (see PIR:OGB07) and V11b (see PIR:OGB07).
C;Function:
A;Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecules of reduced cytochrome c to oxidized cytochrome c, producing two molecules of water and lowering the concentration of cytochrome c.
A;Pathway: oxidative phosphorylation; respiratory chain
A;Note: the role of chain V1a is not clear
C;Superfamily: cytochrome-c oxidase, subunit V1a
C;Keywords: cardiac muscle; electron transfer; heart; membrane-associated complex; mitochondrial protein
F;1-12/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;13-96/Product: cytochrome-c oxidase chain V1a, cardiac #status experimental <MAT>
F;13-24/Domain: mitochondrial matrix #status experimental <MM1>
F;25-49/Domain: transmembrane helix #status experimental <TR01>
F;50-96/Domain: intracisternal #status experimental <ITC1>

Query Match 60.3%; Score 47; DB 1; Length 97;
Best Local Similarity 70.0%; Pred. No. 0.98;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGG 10
||| |||||
Db 43 CTLSWHLHSG 52

RESULT 2
S57895
hypothetical protein 1 - Porphyromonas gingivalis
C;Species: Porphyromonas gingivalis
C;Date: 28-Oct-1995 #sequence_revision 01-Dec-1995 #text_change 17-Mar-2000
C;Accession: S57895
R;Maley, J.; Roberts, I.S.
FEMS Microbiol. Lett. 123, 219-224, 1994
A;Title: Characterisation of IS1126 from Porphyromonas gingivalis W83: a new member of the IS1126 family
A;Reference number: S57895; MUID:95080614; PMID:7988893
A;Accession: S57895
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <MAL>
A;Cross-references: UNIPARC:UPI00001787AF; EMBL:X77924
A;Note: the sequence shown follows the authors' translation displayed in fig. 1
C;Superfamily: transposase IS5

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 7.8;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
:|:| |||||
Db 317 SIRRWFHGGRC 327

RESULT 3
B84228
hypothetical protein Vng0705c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84228
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonka, J.; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <STO>
A;Cross-references: UNIPROT:Q9HRG8; UNIPARC:UPI00000636E5; GB:AE004437; NID:G10580288; P;C;Genetics:
A;Gene: VNG0705C

RESULT 7	
Tl1335	
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Corvus frugilegus mitochondrion	
C;Species: mitochondrion Corvus frugilegus	
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004	
C;Accession: Tl1335	
R;Harlid, A.; Arnason, U.	
Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999	
A;Title: Analyses of mitochondrial DNA nest ratite birds within the Neognathae-supportin	
A;Reference number: Z17262	
A;Accession: Tl1335	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: DNA	
A;Residues: 1-380 <HAR>	
A;Cross-references: UNIPROT:O79386; UNIPARC:UPI0000091E07; EMBL:Y18522; PIDN:CAA77206.1	
C;Genetics:	
A;Genome: mitochondrion	
A;Note: cytb	
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol	
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos	
F;12-340/Domain: cytochrome b homology <CYB>	
F;12-210/Domain: cytochrome b6 homology <CB6>	
F;222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>	
F;84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted	
F;98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted	
Query Match 53.8%; Score 42; DB 2; Length 380;	
Best Local Similarity 80.0%; Pred. No. 26;	
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY 2 TLREWLHGGF 11	
DB 160 TLVEWLWGGF 169	
RESULT 8	
D72073	
conserved hypothetical protein CP0274 [imported] - Chlamydophila pneumoniae (strains CWL	
N;Alternate names: ct383 hypothetical protein	
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae	
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004	
C;Accession: D72073; E81594	
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;	
Nature Genet. 21, 385-389, 1999	
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.	
A;Reference number: A72000; MUID:99206606; PMID:10192388	
A;Accession: D72073	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-218 <ARN>	
A;Cross-references: UNIPROT:Q9Z871; UNIPARC:UPI00000D408B; GB:AE001633; GB:AE001363; NID	
A;Experimental source: strain CWL029	
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,	
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,	
Nucleic Acids Res. 28, 1397-1406, 2000	
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.	
A;Reference number: A81500; MUID:20150255; PMID:10684935	
A;Accession: E81594	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-218 <REA>	
A;Cross-references: UNIPARC:UPI00000D408B; GB:AE002188; GB:AE002161; NID:g7189198; PIDN:	
A;Experimental source: strain AR39, HL cells	
C;Genetics:	
A;Gene: CPn0480; CP0274	
Query Match 52.6%; Score 41; DB 2; Length 218;	
Best Local Similarity 41.7%; Pred. No. 21;	
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;	
QY 1 CTLREWLHGGFC 12	
:: :: :: :: :: ::	

Db 104 CSIVSWVFGGLC 115	
RESULT 9	
D86550	
CT383 hypothetical protein [imported] - Chlamydophila pneumoniae (strain J138)	
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae	
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	
C;Accession: D86550	
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is	
Nucleic Acids Res. 28, 2311-2314, 2000	
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.	
A;Reference number: A86491; MUID:20330349; PMID:10871362	
A;Accession: D86550	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-218 <STO>	
A;Cross-references: UNIPROT:Q9Z871; UNIPARC:UPI00000D408B; GB:BA000008; NID:g8978850; PI	
A;Experimental source: strain J138	
C;Genetics:	
A;Gene: CPj0480	
Query Match 52.6%; Score 41; DB 2; Length 218;	
Best Local Similarity 41.7%; Pred. No. 21;	
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;	
QY 1 CTLREWLHGGFC 12	
:: :: :: :: :: ::	
Db 104 CSIVSWVFGGLC 115	
RESULT 10	
E96566	
F6D8.20 [imported] - Arabidopsis thaliana	
C;Species: Arabidopsis thaliana (mouse-ear cress)	
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004	
C;Accession: E96566	
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,	
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;	
ansen, N.F.; Hughes, B.; Huizar, L.	
Nature 408, 816-820, 2000	
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.	
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,	
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,	
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
A;Reference number: A86141; MUID:21016719; PMID:11130712	
A;Accession: E96566	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-309 <STO>	
A;Cross-references: UNIPROT:Q9SSR0; UNIPARC:UPI00000A4579; GB:AE005173; NID:g5903047; PI.	
C;Genetics:	
A;Gene: F6D8.20	
A;Map position: 1	
C;Superfamily: AARA protein	
Query Match 52.6%; Score 41; DB 2; Length 309;	
Best Local Similarity 100.0%; Pred. No. 30;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 6 WLHGGF 11	
Db 114 WLHGGF 119	
RESULT 11	
A81958	
probable permease NMA0414 [imported] - Neisseria meningitidis (strain 22491 serogroup A)	
C;Species: Neisseria meningitidis	
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004	
C;Accession: A81958	

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-530 <PAR>
A;Cross-references: UNIPROT:Q9JWE3; UNIPARC:UPI00000C497F; GB:AL162753; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0414

Query Match 52.6%; Score 41; DB 2; Length 530;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 3 LREWHLGGGFC 12
|| || || ||
Db 195 LRPWLAGGVC 204

RESULT 12
E81015
ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: E81015
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: E81015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <TET>
A;Cross-references: UNIPROT:Q9JXI9; UNIPARC:UPI00000C4849; GB:AE002552; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2026

Query Match 52.6%; Score 41; DB 2; Length 531;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LREWHLGGGFC 12
|| || || ||
Db 196 LRPWLAGGVC 205

RESULT 13
F82984
hypothetical protein PA5299 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F82984
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F82984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <STO>
A;Cross-references: UNIPROT:Q9HTQ5; UNIPARC:UPI00000C5F94; GB:AE004942; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:

A;Gene: PA5299

Query Match 52.6%; Score 41; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WLHGGF 11
||||||
Db 329 WLHGGF 334

RESULT 14
H86250
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86250
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1025 <STO>
A;Cross-references: UNIPROT:Q9SAA5; UNIPARC:UPI0000048234; GB:AE005172; NID:g4835785; PI
C;Genetics:
A;Map position: 1

Query Match 52.6%; Score 41; DB 2; Length 1025;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WLHGGF 11
||||||
Db 332 WLHGGF 337

RESULT 15
A23443
Pyrimidine synthesis multifunctional protein CAD - golden hamster
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase (C
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 29-Aug-1987 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C;Accession: A38653; A35432; PS0159; A23443; A30794; A34803; I48154
R;Bein, K.; Simmer, J.P.; Evans, D.R.
J. Biol. Chem. 266, 3791-3799, 1991
A;Title: Molecular cloning of a cDNA encoding the amino end of the mammalian multifunct
A;Reference number: A38653; MUID:91139675; PMID:1671675
A;Accession: A38653
A;Molecule type: mRNA
A;Residues: 1-169 <BEI>
A;Cross-references: UNIPROT:P08955; UNIPARC:UPI0000170783; GB:M60078; NID:g191338; PIDN
R;Simmer, J.P.; Kelly, R.E.; Rinker Jr., A.G.; Scully, J.L.; Evans, D.R.
J. Biol. Chem. 265, 10395-10402, 1990
A;Title: Mammalian carbamyl phosphate synthetase (CPS). cDNA sequence and evolution of
A;Reference number: A35432; MUID:90285162; PMID:1972379
A;Accession: A35432
A;Molecule type: mRNA
A;Residues: 156-1455 <SIM>
A;Cross-references: UNIPARC:UPI0000170780; GB:J05503; NID:g191332; PIDN:AAA37062.1; PID
R;Williams, N.K.; Simpson, R.J.; Moritz, R.L.; Peide, Y.; Crofts, L.; Minasian, E.; Lea
Gene 94, 283-288, 1990
A;Title: Location of the dihydroorotase domain within trifunctional hamster dihydroorot
A;Reference number: PS0159; MUID:91078651; PMID:1979549
A;Accession: PS0159

A;Molecule type: mRNA
A;Residues: 1403-2110 <WIL>
A;Cross-references: UNIPARC:UPI00000E5DE8; GB:M33702; NID:g191172; PIDN:AAA37009.1.; PID:
R;Shigesada, K.; Stark, G.R.; Maley, J.A.; Niswander, L.A.; Davidson, J.N.
Mol. Cell. Biol. 5, 1735-1742, 1985
A;Title: Construction of a cDNA to the hamster CAD gene and its application toward defin
A;Reference number: A23443; MUID:85267690; PMID:2862577
A;Accession: A23443
A;Molecule type: mRNA
A;Residues: 2074-2225 <SHI>
A;Cross-references: UNIPARC:UPI00001707A4; GB:M11242; NID:g191330; PIDN:AAA37061.1.; PID:
R;Maley, J.A.; Davidson, J.N.
Biochem. Biophys. Res. Commun. 154, 1047-1053, 1988
A;Title: Identification of the junction between the glutamine amidotransferase and carba
A;Reference number: A30794; MUID:88309082; PMID:2900634
A;Accession: A30794
A;Molecule type: mRNA
A;Residues: 246-464 <MAL>
A;Cross-references: UNIPARC:UPI0000172F70; GB:M21927
R;Simmer, J.P.; Kelly, R.E.; Rinker Jr., A.G.; Zimmermann, B.H.; Scully, J.L.; Kim, H.;
Proc. Natl. Acad. Sci. U.S.A. 87, 174-178, 1990
A;Title: Mammalian dihydroorotase: nucleotide sequence, peptide sequences, and evolution
A;Reference number: A34803; MUID:90115834; PMID:1967494
A;Accession: A34803
A;Molecule type: mRNA
A;Residues: 1391-1870 <SI2>
A;Cross-references: UNIPARC:UPI0000170787; GB:M28866; NID:g191363; PIDN:AAA37073.1.; PID:
A;Note: parts of this sequence were confirmed by peptide sequencing
R;Farnham, P.J.; Kollmar, R.
Cell Growth Differ. 1, 179-189, 1990
A;Title: Characterization of the 5' end of the growth-regulated Syrian hamster CAD gene.
A;Reference number: I48154; MUID:91190717; PMID:1982061
A;Accession: I48154
A;Status: translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-73 <RES>
A;Cross-references: UNIPARC:UPI0000000486; GB:M31621; NID:g191326; PIDN:AAA37060.1.; PID:
C;Genetics:
A;Gene: CAD
A;Map position: B9 short arm
A;Introns: 28/1
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C;Keywords: hydrolase; ligase; methyltransferase; multifunctional enzyme; phosphoprotein
F;4-1442/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;4-354/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolog
F;178-354/Domain: trpG homology <TRG>
F;395-1439/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain hom
F;395-845/Domain: biotin carboxylase homology <BC1>
F;934-1380/Domain: biotin carboxylase homology <BC2>
F;1457-1801/Domain: Bacillus dihydroorotase homology <DHO>
F;1924-2222/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>
F;252/Active site: Cys #status predicted

Query Match 52.6%; Score 41; DB 1; Length 2225;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1 CTLREWL--HG 9
||| ||| ||
Db 102 CTLHSLQQHG 112

RESULT 16
H83343

probable non-ribosomal peptide synthetase PA2424 [imported] - Pseudomonas aeruginosa (sb
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83343
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83343
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4342 <STO>
A;Cross-references: UNIPROT:Q9I157; UNIPARC:UPI0000110218; GB:AE004669; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2424
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;60-553/Domain: acetate-CoA ligase homology <ACL1>
F;584-652/Domain: acyl carrier protein homology <ACP1>
F;1174-1622/Domain: acetate-CoA ligase homology <ACL2>
F;1637-1705/Domain: acyl carrier protein homology <ACP2>
F;2232-2689/Domain: acetate-CoA ligase homology <ACL3>
F;2706-2773/Domain: acyl carrier protein homology <ACP3>
F;3759-4230/Domain: acetate-CoA ligase homology <ACL4>
F;4248-4316/Domain: acyl carrier protein homology <ACP4>
F;1669,2738,4280/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 52.6%; Score 41; DB 2; Length 4342;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLH 8
||:||||:
Db 3559 CTVREWLN 3566

RESULT 17
A38096

perlecan precursor - human
N;Alternate names: basement membrane heparan sulfate proteoglycan; heparan sulfate prote
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 12-Jul-2004
C;Accession: A38096; S19256; S77946; A41059; A40306; B33625; A33625; A41736
R;Murdoch, A.D.; Dodge, G.R.; Cohen, I.; Tuan, R.S.; Iozzo, R.V.
J. Biol. Chem. 267, 8544-8557, 1992
A;Title: Primary structure of the human heparan sulfate proteoglycan from basement membr
tor, laminin, neural cell adhesion molecules, and epidermal growth factor.
A;Reference number: A38096; MUID:92235084; PMID:1569102
A;Accession: A38096
A;Molecule type: mRNA
A;Residues: 1-4391 <MUR>
A;Cross-references: UNIPROT:P98160; UNIPARC:UPI0000168756; GB:M85289; NID:g184426; PIDN:.
R;Kallunki, P.; Tryggvason, K.
J. Cell Biol. 116, 559-571, 1992
A;Title: Human basement membrane heparan sulfate proteoglycan core protein: a 467-kD pro
ell adhesion molecules, and epidermal growth factor.

A;Reference number: A41736; MUID:92112994; PMID:1730768
A;Accession: S19256
A;Molecule type: mRNA
A;Residues: 1-57,'D',59-434,'A',436,'FL',438-449,'Q',451-502,'A',503-792,'K',794-908,'R'
71-2979,'H',2981-2994,'G',2996-3167,'T',3169-3240,'R',3242-3426,'R',3428-3631,'Q',3633-3-
A;Cross-references: UNIPARC:UPI000017CF3E; EMBL:X62515
R;Tryggvason, K.
submitted to the EMBL Data Library, October 1991
A;Reference number: S77946
A;Accession: S77946
A;Molecule type: mRNA
A;Residues: 1-57,'D',59-434,'A',436,'FL',438-449,'Q',451-502,'A',503-792,'K',794-908,'R'
71-2979,'H',2981-2994,'G',2996-3167,'T',3169-3240,'R',3242-3426,'R',3428-3631,'Q',3633-4-
A;Cross-references: UNIPARC:UPI000016866C; EMBL:X62515; NID:g29469; PIDN:CAA44373.1.; PID
R;Kallunki, P.; Eddy, R.L.; Byers, M.G.; Kestilae, M.; Shows, T.B.; Tryggvason, K.
Genomics 11, 389-396, 1991

A;Title: Cloning of human heparan sulfate proteoglycan core protein, assignment of the g
A;Reference number: A41059; MUID:92120660; PMID:1685141
A;Accession: A41059
A;Molecule type: mRNA
A;Residues: 'RT',892-908,'R',910-1101,'L',1103-1132,'L',1134-1221,'L',1223-1397 <KA2>
A;Residues: 'RT',892-908,'R',910-1101,'L',1103-1132,'L',1134-1221,'L',1223-1397 <KA2>
A;Cross-references: UNIPARC:UPI000017C2E3; GB:S76436; NID:g243370; PIDN:AAB21121.1.; PID:.
R;Dodge, G.R.; Kovalszky, I.; Chu, M.L.; Hassell, J.R.; McBride, O.W.; Yi, H.F.; Iozzo,

Genomics 10, 673-680, 1991
A;Title: Heparan sulfate proteoglycan of human colon: partial molecular cloning, cellular
A;Reference number: A40306; MUID:91365376; PMID:1679749
A;Accession: A40306
A;Molecule type: mRNA
A;Residues: 1018-1405, 'G', 1407-1409, 'G', 1411-1465 <DOD>
A;Cross-references: UNIPARC:UPI000017C2E4; GB:M64283; NID:g184424; PIDN:AAA52699.1; PID:
R;Heremans, A.; van der Schueren, B.; De Cock, B.; Paulsson, M.; Cassiman, J.J.; van den
J. Cell Biol. 109, 3199-3211, 1989
A;Title: Matrix-associated heparan sulfate proteoglycan: core protein-specific monoclonal
anes.
A;Reference number: A33625; MUID:90078352; PMID:2687294
A;Accession: B33625
A;Molecule type: protein
A;Residues: 1379-1384, 'X', 1386-1388, 'X', 1390-1398 <HE2>
A;Cross-references: UNIPARC:UPI000017C2E5
A;Accession: A33625
A;Molecule type: protein
A;Residues: 2166-2171, 'X', 2173-2175, 'X', 2177-2185 <HE3>
A;Cross-references: UNIPARC:UPI000017C2E5
A;Note: peptide potentially matches four different regions of sequence shown
C;Genetics:
A;Gene: GDB:HSPG2
A;Cross-references: GDB:126372; OMIM:142461
A;Map position: lp36.1-lp36.1
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein; heparan sulfate; transmembra
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-4391/Product: perlecan #status predicted <MAT>
F;22-193/Domain: I <DOM1>
F;194-530/Domain: II <DOM2>
F;199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;531-1676/Domain: III <DOM3>
F;1159-1206/Domain: laminin-type EGF-like homology <LEG>
F;1563-1610/Domain: laminin-type EGF-like homology <EG7>
F;1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F;1677-3686/Domain: IV <DOM4>
F;2007-2034/Domain: transmembrane #status predicted <TRM>
F;3687-4391/Domain: V <DOM5>
F;3845-3880/Domain: EGF homology <EGF1>
F;3888-3921/Domain: EGF homology <EGF>
F;3953-4106/Domain: laminin G repeat homology <LG2>
F;4147-4175/Domain: EGF homology <EGF2>
F;4149-4151/Region: motor neuron attachment (L-R-E) motif
F;4299-4301/Region: motor neuron attachment (L-R-E) motif
F;65, 71, 76/Binding site: heparan sulfate (Ser) (covalent) #status predicted
F;89, 554, 1755, 2121, 3072, 3105, 3279, 3780, 3836, 4068/Binding site: carbohydrate (Asn) (coval
F;2995, 3933, 4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 51.9%; Score 40.5; DB 2; Length 4391;
Best Local Similarity 69.2%; Pred. No. 5.3e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CTLRE-WLHGGFC 12
Db 4147 CQLREPCVLHGGTC 4159

RESULT 18
B87354
hypothetical protein CC0845 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87354
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87354

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-117 <STO>
A;Cross-references: UNIPROT:Q9A9W6; UNIPARC:UPI00000C71B8; GB:AE005673; NID:g13422100; P
C;Genetics:
A;Gene: CC0845
Query Match 51.3%; Score 40; DB 2; Length 117;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
Db 38 CVLGHWLHG 46

RESULT 19
S25970
hypothetical protein 139 - liverwort (Marchantia polymorpha) mitochondrion
C;Species: mitochondrion Marchantia polymorpha
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 09-Jul-2004
C;Accession: S25970
R;Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.; Kan
J. Mol. Biol. 223, 1-7, 1992
A;Title: Gene organization deduced from the complete sequence of liverwort Marchantia po
A;Reference number: S25941; MUID:92114051; PMID:1731062
A;Accession: S25970
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-139 <ODA>
A;Cross-references: UNIPROT:P38470; UNIPARC:UPI000013B8D1; EMBL:M68929; NID:g786182; PFI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
C;Genetics:
A;Genome: mitochondrion
C;Superfamily: liverwort mitochondrion hypothetical protein 139
C;Keywords: mitochondrion

Query Match 51.3%; Score 40; DB 2; Length 139;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 REWLHGGFC 12
Db 78 REWFLGGWC 86

RESULT 20
G84305
hypothetical protein muty [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84305
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <STO>
A;Cross-references: UNIPROT:Q9HPQ6; UNIPARC:UPI00000638EB; GB:AE004437; NID:g10581009;
C;Genetics:
A;Gene: muty

Query Match 51.3%; Score 40; DB 2; Length 312;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLH 8
| | | | | |

Db 215 CPLREWCH 222

RESULT 21
S39007

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - guinea pig mitochondrion
C;Species: mitochondrion Cavia porcellus (guinea pig)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: S39007
R;Ma, D.P.; Zharkikh, A.; Graur, D.; VandeBerg, J.L.; Li, W.H.
J. Mol. Evol. 36, 327-334, 1993
A;Title: Structure and evolution of opossum, guinea pig, and porcupine cytochrome b gene
A;Reference number: S33572; MUID:93301932; PMID:8315653
A;Accession: S39007
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-378 <MAD>
A;Cross-references: UNIPROT:Q9TEF9; UNIPARC:UPI0000174C8F
C;Genetics:
A;Gene: cob
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;10-338/Domain: cytochrome b homology <CBH>
F;10-208/Domain: cytochrome b6 homology <CB6>
F;35-51/Domain: transmembrane #status predicted <TM1>
F;80-98/Domain: transmembrane #status predicted <TM2>
F;116-132/Domain: transmembrane #status predicted <TM3>
F;177-199/Domain: transmembrane #status predicted <TM4>
F;220-338/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-244/Domain: transmembrane #status predicted <TM5>
F;287-303/Domain: transmembrane #status predicted <TM6>
F;322-342/Domain: transmembrane #status predicted <TM7>
F;352-368/Domain: transmembrane #status predicted <TM8>
F;82,181/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;96,195/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 378;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||: |||
Db 158 TLVEWIWGGF 167

RESULT 22
S17406

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Arabian camel mitochondrion
C;Species: mitochondrion Camelus dromedarius (Arabian camel)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17406
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17406
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24952; UNIPARC:UPI000012895E; EMBL:X56281; NID:g12854; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-209/Domain: cytochrome b6 homology <CB6>

Query Match 51.3%; Score 40; DB 2; Length 378;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||: |||
Db 158 TLVEWIWGGF 167

RESULT 22
S17406

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Arabian camel mitochondrion
C;Species: mitochondrion Camelus dromedarius (Arabian camel)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17406
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17406
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24952; UNIPARC:UPI000012895E; EMBL:X56281; NID:g12854; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosme phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>

F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||: |||
Db 159 TLVEWIWGGF 168

RESULT 23
S17408

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - black rhinoceros mitochond
C;Species: mitochondrion Diceror bicornis (black rhinoceros)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17408
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17408
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24954; UNIPARC:UPI00001289A5; EMBL:X56283; NID:g12903; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||: |||
Db 159 TLVEWIWGGF 168

RESULT 24
S17410

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Grevy's zebra mitochondrio
C;Species: mitochondrion Equus grevyi (Grevy's zebra)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004

C;Accession: S17410
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17410
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24956; UNIPARC:UPI00001289C7; EMBL:X56282; NID:g12948; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 25
S17415
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - bridled dolphin mitochondr
C;Species: mitochondrion Stenella attenuata (bridled dolphin)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17415
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17415
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24961; UNIPARC:UPI000016D629; EMBL:X56294; NID:g13495; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 26
S17417
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b, isolate 1B - pantropical sp
C;Species: mitochondrion Stenella longirostris (pantropical spinner dolphin)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17417; S17416
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17417
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24962; UNIPARC:UPI00000001E0; EMBL:X56293; NID:g13628; PIDN
A;Experimental source: isolate 1B
A;Accession: S17416
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-59,'T',61-97,'M',99-265,'P',267-299,'I',301-326,'V',328-379 <IR2>
A;Cross-references: UNIPARC:UPI0000128B12; EMBL:X56292; NID:g13626; PIDN:CAA39739.1; PIL
A;Experimental source: isolate 1A
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase compl
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rel
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 27
S33572
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - southern African porcupin.

C;Species: mitochondrion Hystrix africaeaustralis (southern African porcupine)
C;Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S33572
R;Ma, D.P.; Zharkikh, A.; Graur, D.; VandeBerg, J.L.; Li, W.H.
J. Mol. Evol. 36, 327-334, 1993
A;Title: Structure and evolution of opossum, guinea pig, and porcupine cytochrome b gene
A;Reference number: S33572; MUID:93301932; PMID:8315653
A;Accession: S33572
A;Molecule type: DNA
A;Residues: 1-379 <MAD>
A;Cross-references: UNIPROT:Q04910; UNIPARC:UPI000016D69D; EMBL:X70674; NID:gl4012; PIDN
A;Note: residue 1 and the corresponding nucleotide sequence are not shown
C;Genetics:
A;Gene: cob
A;Genome: mitochondrion
A;Genetic code: SGCl
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 28
S41832
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - blue whale mitochondrion
C;Species: mitochondrion Balaenoptera musculus (blue whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S41832
R;Arnason, U.; Gullberg, A.
J. Mol. Evol. 37, 312-322, 1993
A;Title: Comparison between the complete mtDNA sequences of the blue and the fin whale,
A;Reference number: S41820; MUID:94141932; PMID:8308901
A;Accession: S41832
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41285; UNIPARC:UPI0000128949; EMBL:X72204; NID:g414126; PID
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCl
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 29
S43261
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - minke whale mitochondrion
C;Species: mitochondrion Balaenoptera acutorostrata (minke whale, lesser rorqual)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43261
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43261
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41280; UNIPARC:UPI0000128944; EMBL:X75753; NID:g457761; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCl
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 30
S43262
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaenoptera bonaerensis m
C;Species: mitochondrion Balaenoptera bonaerensis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43262
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43262
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 29
S43261
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - minke whale mitochondrion
C;Species: mitochondrion Balaenoptera acutorostrata (minke whale, lesser rorqual)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43261
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43261
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41280; UNIPARC:UPI0000128944; EMBL:X75753; NID:g457761; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCl
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 30
S43262
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaenoptera bonaerensis m
C;Species: mitochondrion Balaenoptera bonaerensis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43262
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43262
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41281; UNIPARC:UPI0000128945; EMBL:X75581; NID:G457762; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted
Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 TLREWLHGPF 11
Db 159 TLVEWIWGGF 168
RESULT 31
S43263
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - sei whale mitochondrion
C;Species: mitochondrion Balaenoptera borealis (sei whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43263
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43263
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41282; UNIPARC:UPI0000128946; EMBL:X75582; NID:G457763; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted
Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 TLREWLHGPF 11

Db 159 TLVEWIWGGF 168
RESULT 32
S43264
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Bryde's whale mitochondrio
C;Species: mitochondrion Balaenoptera edeni (Bryde's whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43264
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43264
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41283; UNIPARC:UPI0000128947; EMBL:X75583; NID:G457766; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) #status predicted
Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 TLREWLHGPF 11
Db 159 TLVEWIWGGF 168
RESULT 33
S43265
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena glacialis mitoch
C;Species: mitochondrion Balaena glacialis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43265
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43265
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41284; UNIPARC:UPI0000171CE6; EMBL:X75587
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>

F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||::|||
Db 159 TLVEWIWGGF 168

RESULT 34
S43266
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena mysticetus mitochondrion
C;Species: mitochondrion Balaena mysticetus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43266
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994

A;Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A;Reference number: S43261; MUID:94150700; PMID:8107866

A;Accession: S43266
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41286; UNIPARC:UPI000012894A; EMBL:X75588; NID:g457770; PID:107866
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:

A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m

in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||::|||
Db 159 TLVEWIWGGF 168

RESULT 35
S43267
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Caperea marginata mitochondrion
C;Species: mitochondrion Caperea marginata
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43267
R;Arnason, U.; Gullberg, A.

Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43267
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41287; UNIPARC:UPI0000128967; EMBL:X75586; NID:g457773; PID:107866
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:

A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m

in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||::|||
Db 159 TLVEWIWGGF 168

RESULT 36
S43269

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - humpback whale mitochondrion
C;Species: mitochondrion Megaptera novaeangliae (humpback whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43269
R;Arnason, U.; Gullberg, A.

Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43269

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41289; UNIPARC:UPI0000128A3A; EMBL:X75584; NID:g457794; PID:107866
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:

A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m

in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||::|||
Db 159 TLVEWIWGGF 168

RESULT 35
S43267
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Caperea marginata mitochondrion
C;Species: mitochondrion Caperea marginata
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43267
R;Arnason, U.; Gullberg, A.

Query Match		51.3%;	Score 40;	DB 1;	Length 379;
Best Local Similarity		70.0%;	Pred. No. 54;		
Matches	7;	Conservative	1;	Mismatches	2; Indels 0; Gaps 0;
QY	2 TLREWLHGGF 11				
	::				
Db	159 TLVEWIWGGF 168				
RESULT 37					
E58889					
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - duckbill platypus mitochondrion					
C;Species: mitochondrion Ornithorhynchus anatinus (duckbill platypus)					
C;Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004					
C;Accession: E58889					
R;Janke, A.; Gemmell, N.J.; Feldmaier-Fuchs, G.; von Haeseler, A.; Paabo, S.					
J. Mol. Evol. 42, 153-159, 1996					
A;Title: The mitochondrial genome of a monotreme--the platypus (Ornithorhynchus anatinus)					
A;Reference number: A58888; MUID:97077300; PMID:8919867					
A;Accession: E58889					
A;Status: nucleic acid sequence not shown; translation not shown; not compared with cons					
A;Molecule type: DNA					
A;Residues: 1-379 <JAN>					
A;Cross-references: UNIPROT:Q36461; UNIPARC:UPI0000128A72; GB:X83427; NID:g1469249; PIDN					
A;Note: submitted to GenBank/EMBL/DDBJ December, 1994					
C;Genetics:					
A;Map position: FOR14186-15325					
A;Genome: mitochondrion					
A;Genetic code: SGC1					
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol					
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;					
F;11-339/Domain: cytochrome b homology <CBH>					
F;11-209/Domain: cytochrome b6 homology <CB6>					
F;36-52/Domain: transmembrane #status predicted <TM1>					
F;81-99/Domain: transmembrane #status predicted <TM2>					
F;117-133/Domain: transmembrane #status predicted <TM3>					
F;142-146/Region: ubiquinone binding #status predicted					
F;178-200/Domain: transmembrane #status predicted <TM4>					
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>					
F;229-245/Domain: transmembrane #status predicted <TM5>					
F;288-304/Domain: transmembrane #status predicted <TM6>					
F;323-343/Domain: transmembrane #status predicted <TM7>					
F;353-369/Domain: transmembrane #status predicted <TM8>					
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted					
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted					
Query Match		51.3%;	Score 40;	DB 2;	Length 379;
Best Local Similarity		70.0%;	Pred. No. 54;		
Matches	7;	Conservative	1;	Mismatches	2; Indels 0; Gaps 0;
QY	2 TLREWLHGGF 11				
	::				
Db	159 TLVEWIWGGF 168				
RESULT 38					
I48133					
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl					
C;Species: mitochondrion Dactylomys boliviensis (American spiny rat)					
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004					
C;Accession: I48133					
R;da Silva, M.N.F.; Patton, J.L.					
Mol. Phylogenet. Evol. 2, 243-255, 1993					
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyid rodents					
A;Reference number: A49605; MUID:94184505; PMID:8136924					
A;Accession: I48133					
A;Status: preliminary; translated from GB/EMBL/DDBJ					
A;Molecule type: DNA					
A;Residues: 1-379 <RES>					
A;Cross-references: UNIPROT:Q34301; UNIPARC:UPI0000095C92; GB:L233339; NID:g996082; PIDN					
C;Genetics:					
A;Genome: mitochondrion					

A;Genetic code: SGC1	
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol	
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;	
F;11-339/Domain: cytochrome b homology <CBH>	
F;11-209/Domain: cytochrome b6 homology <CB6>	
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>	
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted	
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicte	
Query Match	
Best Local Similarity	
Matches	7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY	2 TLREWLHGGF 11
	::
Db	159 TLVEWIWGGF 168
RESULT 39	
I48132	
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl	
C;Species: mitochondrion Dactylomys dactylinus (American spiny rat)	
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004	
C;Accession: I48132	
R;da Silva, M.N.F.; Patton, J.L.	
Mol. Phylogenet. Evol. 2, 243-255, 1993	
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyid rodents	
A;Reference number: A49605; MUID:94184505; PMID:8136924	
A;Accession: I48132	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: DNA	
A;Residues: 1-379 <RES>	
A;Cross-references: UNIPROT:Q34306; UNIPARC:UPI0000092E09; GB:L23335; NID:g995850; PIDN:	
C;Genetics:	
A;Genome: mitochondrion	
A;Genetic code: SGC1	
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol	
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;	
F;11-339/Domain: cytochrome b homology <CBH>	
F;11-209/Domain: cytochrome b6 homology <CB6>	
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>	
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted	
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predict	
Query Match	
Best Local Similarity	
Matches	7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY	2 TLREWLHGGF 11
	::
Db	159 TLVEWIWGGF 168
RESULT 40	
I48180	
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Isoth	
C;Species: mitochondrion Isothrix bistriatus (American spiny rat)	
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004	
C;Accession: I48180	
R;da Silva, M.N.F.; Patton, J.L.	
Mol. Phylogenet. Evol. 2, 243-255, 1993	
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyid rodent	
A;Reference number: A49605; MUID:94184505; PMID:8136924	
A;Accession: I48180	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: DNA	
A;Residues: 1-379 <RES>	
A;Cross-references: UNIPROT:Q34814; UNIPARC:UPI0000094264; GB:L233355; NID:g995846; PIDN:	
C;Genetics:	
A;Genome: mitochondrion	
A;Genetic code: SGC1	
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquino	
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion	

F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 41
E58851
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - finback whale mitochondrion
C;Species: mitochondrion Balaenoptera physalus (finback whale, common rorqual)
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: E58851
R;Arnason, U.; Gullberg, A.; Widegren, B.
J. Mol. Evol. 33, 556-568, 1991
A;Title: The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Bal
A;Reference number: A58850; MUID:92139449; PMID:1779436
A;Accession: E58851
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P24950; UNIPARC:UPI000012894C; GB:X61145; NID:g12772; PIDN:C
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in

F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 42
T11375
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - donkey mitochondrion
C;Species: mitochondrion Equus asinus (donkey)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11375
R;Xu, X.; Gullberg, A.; Arnason, U.
J. Mol. Evol. 43, 438-463, 1996
A;Title: The complete mitochondrial (mtDNA) of the donkey and mtDNA comparisons among fo
A;Reference number: Z17265; MUID:97032591; PMID:8875857
A;Accession: T11375
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-379 <XUX>
A;Cross-references: UNIPROT:P92487; UNIPARC:UPI00001289C6; EMBL:X97337; NID:g1805746; PI

A;Experimental source: kidney
C;Genetics:
A;Genome: mitochondrion
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 43
T11869
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - horse mitochondrion
C;Species: mitochondrion Equus caballus (domestic horse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11869
R;Xu, X.; Arnason, U.
Gene 148, 357-362, 1994
A;Title: The complete mitochondrial DNA sequence of the horse, Equus caballus: Extensive
A;Reference number: Z17369; MUID:95047450; PMID:7958969
A;Accession: T11869
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-379 <XUX>
A;Cross-references: UNIPROT:P48665; UNIPARC:UPI00001289F7; EMBL:X79547; NID:g577571; PID
C;Genetics:
A;Genome: mitochondrion
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 44
T11492
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rabbit mitochondrion
C;Species: mitochondrion Oryctolagus cuniculus (domestic rabbit)
C;Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T11492; S13823
R;Gissi, C.; Gullberg, A.; Arnason, U.
Genomics 50, 161-169, 1998
A;Title: The complete mitochondrial DNA sequence of the rabbit, Oryctolagus cuniculus.
A;Reference number: Z17275; MUID:98317530; PMID:9653643
A;Accession: T11492
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-379 <GIS>
A;Cross-references: UNIPROT:P34863; UNIPARC:UPI0000128AC6; EMBL:AJ001588; NID:g3293006;
A;Note: the termination resulting from transcript polyadenylation is shown
R;Mignotte, F.; Gueride, M.; Champagne, A.M.; Mounolou, J.C.
Eur. J. Biochem. 194, 561-571, 1990
A;Title: Direct repeats in the non-coding region of rabbit mitochondrial DNA. Involvement

A;Reference number: S13823; MUID:91099332; PMID:2269281
A;Accession: S13823
A;Molecule type: DNA
A;Residues: 211-379 <EUR>
A;Cross-references: UNIPARC:UPI000016D6B1; EMBL:X54172; NID:g14141; PIDN:CAA38105.1; PID
C;Genetics:
A;Gene: cytb
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b6 homology <CB6>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 45
S58990
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - earthworm (Lumbricus terre
C;Species: mitochondrion Lumbricus terrestris (common earthworm)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C;Accession: S58990
R;Boore, J.L.; Brown, W.M.
Genetics 141, 305-319, 1995
A;Title: Complete sequence of the mitochondrial DNA of the annelid worm Lumbricus terre
A;Reference number: S58985; MUID:96042914; PMID:8536978
A;Accession: S58990
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <BOO>
A;Cross-references: UNIPROT:Q34945; UNIPARC:UPI0000128A2D; EMBL:U24570; NID:g984290; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC4
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;12-340/Domain: cytochrome b homology <CBH>
F;12-210/Domain: cytochrome b6 homology <CB6>
F;222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 160 TLVEWIWGGF 169

RESULT 46
T11401
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - fat dormouse mitochondrion
C;Species: mitochondrion Myoxus glis (fat dormouse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11401
R;Reyes, A.; Pesole, G.; Saccone, C.
Mol. Biol. Evol. 15, 499-505, 1998
A;Title: Complete mitochondrial DNA sequence of the fat dormouse, Glis glis: further evi
A;Reference number: Z17267; MUID:98242079; PMID:9580978
A;Accession: T11401

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-379 <REY>
A;Cross-references: UNIPROT:O63910; UNIPARC:UPI000009684F; EMBL:AJ001562; NID:g3127895;
C;Genetics:
A;Gene: cytb
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b6 homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 47
A53077
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Spermophilus richardsonii
C;Species: mitochondrion Spermophilus richardsonii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C;Accession: A53077
R;Thomas, W.K.; Martin, S.L.
Mol. Phylogenet. Evol. 2, 330-336, 1993
A;Title: A recent origin of marmots.
A;Reference number: A53077; MUID:94326000; PMID:8049781
A;Accession: A53077
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-379 <RES>
A;Cross-references: UNIPROT:P49341; UNIPARC:UPI0000128B0A; GB:S73150; NID:g639944; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predict

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 48
I48134
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Echim
C;Species: mitochondrion Echimy's chrysurus (American spiny rat)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 31-Dec-2004
C;Accession: I48134
R;da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodent
A;Reference number: A49605; MUID:94184505; PMID:8136924
A;Accession: I48134
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A/Residues: 1-379 <RES>
A/Cross-references: UNIPROT:Q34420; UNIPARC:UPI0000096931; GB:L23341; NID:g995848; PIDN:
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Superfamily: cytochrome b homology; cytochrome b6 homology; plastoquinol-plastocyanin
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 49
CBRT
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rat mitochondrion
C/Species: mitochondrion Rattus norvegicus (Norway rat)
C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C/Accession: A00154; S04759
R/Koike, K.; Kobayashi, M.; Yaginuma, K.; Taira, M.; Yoshida, E.; Imai, M.
Gene 20, 177-185, 1982
A/Title: Nucleotide sequence and evolution of the rat mitochondrial cytochrome b gene co
A/Reference number: A00154; MUID:83158755; PMID:6299885
A/Accession: A00154
A/Molecule type: DNA
A/Residues: 1-380 <KOI>
A/Cross-references: UNIPROT:P00159; UNIPARC:UPI000016D61E; GB:J01436; NID:g343168; PIDN:
A/Note: the authors translated the codon ATA for residue 42 as Ile, CAC for residue 54 a
96 as Phe. In another figure, the amino acids at all these positions except position 89,
translation in having 11-Glu
R/Gadaleta, G.; Pepe, G.; De Candia, G.; QuagliarIELlo, C.; Sbisa, E.; Saccone, C.
J. Mol. Evol. 28, 497-516, 1989
A/Title: The complete nucleotide sequence of the Rattus norvegicus mitochondrial genome:
A/Reference number: S04747; MUID:89362487; PMID:2504926
A/Accession: S04759
A/Molecule type: DNA
A/Residues: 1-82,'Q',84-152,'I',154-380 <GAD>
A/Cross-references: UNIPARC:UPI000016390E; EMBL:X14848; NID:g854269; PIDN:CAA32966.1; PI
C/Genetics:
A/Gene: cob
A/Map position: 87-94
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Function:
A/Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A/Pathway: oxidative phosphorylation; respiratory chain
C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 380;

Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

RESULT 50
CBMS
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - mouse mitochondrion
C/Species: mitochondrion Mus musculus (house mouse)
C/Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C/Accession: A00153
R/Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
Cell 26, 167-180, 1981
A/Title: Sequence and gene organization of mouse mitochondrial DNA.
A/Reference number: A00153; MUID:82137051; PMID:7332926
A/Accession: A00153
A/Molecule type: DNA
A/Residues: 1-381 <BIB>
A/Cross-references: UNIPROT:P00158; UNIPARC:UPI00000008FF; GB:J01420; NID:g342520; PIDN:
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Function:
A/Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A/Pathway: oxidative phosphorylation; respiratory chain
C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 381;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIWGGF 168

Search completed: May 12, 2006, 10:51:22
Job time : 32.6641 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 84.9231 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTLREWLHGGFC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	47	60.3	97	1	CX6A2_BOVIN	P07471 bos taurus
2	47	60.3	128	2	Q9GAB0_9BIVA	Q9gab0 pitar morrh
3	46	59.0	128	2	Q9GAB2_MERCA	Q9gab2 mercenaria
4	46	59.0	128	2	Q9GAB1_9BIVA	Q9gab1 calyptogena
5	46	59.0	279	2	Q6H5Y4_ORYSA	Q6h5y4 oryza sativ
6	46	59.0	462	2	Q75D31_ASHGO	Q75d31 ashbya gos
7	46	59.0	767	2	Q4XYI2_PLACH	Q4xyi2 plasmodium
8	46	59.0	1429	2	Q4YVE2_PLABE	Q4yve2 plasmodium
9	46	59.0	2558	2	Q8IKF6_PLAF7	Q8ikf6 plasmodium
10	46	59.0	2584	2	Q7PDT6_PLAYO	Q7pdt6 plasmodium
11	45	57.7	126	2	Q8TJE0_METAC	Q8tje0 methanosarc
12	45	57.7	211	2	Q92NB4_PORGI	Q9znb4 porphyromon
13	45	57.7	211	2	Q7MUZ9_PORGI	Q7muz9 porphyromon
14	45	57.7	361	2	Q9ZA61_PORGI	Q9za61 porphyromon
15	45	57.7	361	2	Q9ZAD0_PORGI	Q9zad0 porphyromon
16	45	57.7	361	2	Q7M7B5_PORGI	Q7m7b5 porphyromon
17	45	57.7	361	2	Q7M7E9_PORGI	Q7m7e9 porphyromon
18	45	57.7	396	2	Q67UT0_ORYSA	Q67ut0 oryza sativ
19	45	57.7	503	2	Q9HRG8_HALSA	Q9hrg8 halobacteri
20	45	57.7	532	1	CBPY_YEAST	P00729 saccharomyc
21	44	56.4	154	2	Q7V2A4_PROMP	Q7v2a4 prochloroco
22	44	56.4	171	2	Q9ZPU0_ARATH	Q9zpu0 arabidopsis
23	44	56.4	539	2	Q6CBE1_YARLI	Q6cbe1 yarrowia li
24	44	56.4	731	2	Q61MS4_CAEBR	Q61ms4 caenorhabdi
25	44	56.4	1540	2	Q628B7_CAEBR	Q628b7 caenorhabdi
26	44	56.4	4856	2	Q61EJ2_CAEBR	Q61ej2 caenorhabdi
27	43	55.1	158	2	Q7U636_SYNXP	Q7u636 synechococc
28	43	55.1	182	2	Q9B598_CHILA	Q9b598 chinchilla
29	43	55.1	328	2	Q5V5F2_HALMA	Q5v5f2 haloarcula
30	43	55.1	380	2	Q9B640_9RODE	Q9b640 eothenomys
31	43	55.1	440	2	Q8WF54_9BIVA	Q8wf54 venerupis (

32	43	55.1	558	2	Q4ICS6_GIBZE	Q4ics6 gibberella
33	43	55.1	583	2	Q8CY51_BIFLO	Q8cy51 bifidobacte
34	43	55.1	605	2	Q59GS6_HUMAN	Q59gs6 homo sapien
35	43	55.1	766	2	Q5FWG1_HUMAN	Q5fwg1 homo sapien
36	43	55.1	791	1	DGKG_HUMAN	p49619 homo sapien
37	43	55.1	884	2	Q53JX1_ORYSA	Q53jx1 oryza sativ
38	42.5	54.5	104	2	Q9D245_MOUSE	Q9d245 m mus muscu
39	42	53.8	58	2	Q9T204_BPLC2	Q9t204 lactococcus
40	42	53.8	112	2	Q85A85_CORMC	Q85a85 corvus macr
41	42	53.8	121	2	Q8BG00_MOUSE	Q8bg00 m mus muscu
42	42	53.8	128	2	P92837_9HYME	P92837 ephedrus pe
43	42	53.8	134	2	Q7YAZ1_9SAUR	Q7yaz1 colobosaura
44	42	53.8	144	2	Q9MMC3_GUIGU	Q9mmc3 guira guira
45	42	53.8	144	2	Q9MMC4_9AVES	Q9mmc4 crotophaga
46	42	53.8	193	2	Q4VTU6_9SAUR	Q4vtu6 gymnodactyl
47	42	53.8	196	2	Q7YB81_9SYLV	Q7yb81 alcippe mor
48	42	53.8	203	2	Q4VTU5_9SAUR	Q4vtu5 gymnodactyl
49	42	53.8	213	2	Q4VTU4_9SAUR	Q4vtu4 gymnodactyl
50	42	53.8	215	2	Q34157_CORMC	Q34157 corvus macr
51	42	53.8	215	2	Q34078_9SYLV	Q34078 cettia diph
52	42	53.8	215	2	Q34901_LEILU	Q34901 leiothrix l
53	42	53.8	221	2	Q8M2E3_9SAUR	Q8m2e3 pachydactyl
54	42	53.8	221	2	Q8M2D8_9SAUR	Q8m2d8 pachydactyl
55	42	53.8	222	2	Q7YB76_9SYLV	Q7yb76 alcippe poi
56	42	53.8	223	2	Q9TGR7_9SAUR	Q9tgr7 aspidosceli
57	42	53.8	226	2	Q7YB78_9SYLV	Q7yb78 alcippe per
58	42	53.8	226	2	Q7YAZ4_9SAUR	Q7yaz4 cordylus na
59	42	53.8	228	2	Q9MHR8_9SAUR	Q9mhr8 hemidactylu
60	42	53.8	233	2	Q7YAZ0_9SAUR	Q7yaz0 leposoma sc
61	42	53.8	233	2	Q7YAV2_9SAUR	Q7yav2 cnemidophor
62	42	53.8	238	2	Q34569_GUIGU	Q34569 guira guira
63	42	53.8	238	2	Q4VTU3_9SAUR	Q4vtu3 gymnodactyl
64	42	53.8	249	2	O21086_9RODE	O21086 peromyscus
65	42	53.8	249	2	O21085_9RODE	O21085 peromyscus
66	42	53.8	252	2	Q7YB80_9SYLV	Q7yb80 alcippe mor
67	42	53.8	253	2	Q7YB77_9SYLV	Q7yb77 alcippe per
68	42	53.8	256	2	Q9TCQ6_9SAUR	Q9tcg6 aspidosceli
69	42	53.8	258	2	Q9TCQ5_9SAUR	Q9tcg5 aspidosceli
70	42	53.8	261	2	Q956Q5_9SAUR	Q956q5 aspidosceli
71	42	53.8	261	2	Q9TGP4_9SAUR	Q9tgp4 aspidosceli
72	42	53.8	263	2	Q9TCQ7_9SAUR	Q9tcg7 aspidosceli
73	42	53.8	263	2	Q9TGP9_9SAUR	Q9tgp9 aspidosceli
74	42	53.8	263	2	Q9TGP8_9SAUR	Q9tgp8 aspidosceli
75	42	53.8	264	2	Q9TGR8_9SAUR	Q9tgr8 aspidosceli
76	42	53.8	266	2	Q9TGR3_9SAUR	Q9tgr3 aspidosceli
77	42	53.8	266	2	Q9TGP9_9SAUR	Q9tgp9 aspidosceli
78	42	53.8	266	2	Q9TGP2_9SAUR	Q9tgp2 aspidosceli
79	42	53.8	266	2	Q9TGP8_9SAUR	Q9tgp8 aspidosceli
80	42	53.8	267	2	Q9TGR4_9SAUR	Q9tgr4 aspidosceli
81	42	53.8	267	2	Q9XKP3_9SAUR	Q9xkp3 aspidosceli
82	42	53.8	267	2	Q9TGP4_9SAUR	Q9tgp4 aspidosceli
83	42	53.8	267	2	Q9XKN6_9SAUR	Q9xkn6 aspidosceli
84	42	53.8	267	2	Q9TGR5_9SAUR	Q9tgr5 aspidosceli
85	42	53.8	268	2	Q9TGR0_9SAUR	Q9tgr0 aspidosceli
86	42	53.8	268	2	Q9XKP1_9SAUR	Q9xkp1 aspidosceli
87	42	53.8	268	2	Q956Q6_9SAUR	Q956q6 aspidosceli
88	42	53.8	268	2	Q9TGP3_9SAUR	Q9tgp3 aspidosceli
89	42	53.8	268	2	Q9TGP6_9SAUR	Q9tgp6 aspidosceli
90	42	53.8	269	2	Q9TGP0_9SAUR	Q9tgp0 aspidosceli
91	42	53.8	269	2	Q9XKN5_9SAUR	Q9xkn5 aspidosceli
92	42	53.8	269	2	Q9TGR6_9SAUR	Q9tgr6 aspidosceli
93	42	53.8	270	2	Q9TGR1_9SAUR	Q9tgr1 aspidosceli
94	42	53.8	270	2	Q6UKX4_9SAUR	Q6ukx4 liolaemus p
95	42	53.8	271	2	Q9TGP7_9SAUR	Q9tgp7 aspidosceli
96	42	53.8	271	2	Q9XKP2_9SAUR	Q9xkp2 aspidosceli
97	42	53.8	271	2	Q9TGP3_9SAUR	Q9tgp3 aspidosceli
98	42	53.8	271	2	Q9XKN9_9SAUR	Q9xkn9 aspidosceli
99	42	53.8	271	2	Q8M4S1_9SYLV	Q8m4s1 alcippe poi
100	42	53.8	271	2	Q9TGP1_9SAUR	Q9tgp1 aspidosceli
101	42	53.8	271	2	Q9XKN8_9SAUR	Q9xkn8 aspidosceli
102	42	53.8	271	2	Q9XKN7_9SAUR	Q9xkn7 aspidosceli
103	42	53.8	271	2	Q9XKP0_9SAUR	Q9xkp0 aspidosceli
104	42	53.8	271	2	Q9TGR2_9SAUR	Q9tgr2 aspidosceli

251	42	53.8	381	1	CYB_PLAGI	Q35459 planigale g	324	41	52.6	379	2	Q9T6I2_LEPSI	Q9t6i2 lepus sinen
252	42	53.8	381	1	CYB_PLAIN	Q35485 planigale i	325	41	52.6	380	1	CYB_THOIS	Q9xnw9 thomasomys
253	42	53.8	381	1	CYB_PLATE	Q35675 planigale t	326	41	52.6	381	1	CYB_ANTIME	Q33782 antechinus
254	42	53.8	381	1	CYB_PSEMO	O03543 pseudantech	327	41	52.6	381	1	CYB_THYCY	Q36012 thylacinus
255	42	53.8	381	1	CYB_PSEWO	Q35695 pseudantech	328	41	52.6	381	2	Q9XP74_SMIVI	Q9xp74 sminthopsis
256	42	53.8	381	1	CYB_SARHA	Q35861 sarcophilus	329	41	52.6	391	2	Q4KAE2_PSEF5	Q4kae2 pseudomonas
257	42	53.8	381	1	CYB_SMIAI	Q9xp89 sminthopsis	330	41	52.6	391	2	Q4SIH1_TETNG	Q4sih1 tetraodon n
258	42	53.8	381	1	CYB_SMIAI	Q9xp88 sminthopsis	331	41	52.6	460	2	Q7PR34_ANOGA	Q7pr34 anopheles g
259	42	53.8	381	1	CYB_SMICR	Q35810 sminthopsis	332	41	52.6	494	2	Q60VG8_CAEBR	Q60vg8 caenorhabdi
260	42	53.8	381	1	CYB_SMIGA	Q9xp83 sminthopsis	333	41	52.6	530	2	Q5F558_NEIG1	Q5f558 neisseria g
261	42	53.8	381	1	CYB_SMIMA	O20433 sminthopsis	334	41	52.6	530	2	Q9JWE3_NEIMA	Q9jwe3 neisseria m
262	42	53.8	381	1	CYB_SMIMU	Q35886 sminthopsis	335	41	52.6	531	2	Q9JX19_NEIMB	Q9jxi9 neisseria m
263	42	53.8	381	2	Q9XP80_SMILE	Q9xp80 sminthopsis	336	41	52.6	542	2	Q8PF22_XANAC	Q8pf22 xanthomonas
264	42	53.8	381	2	Q9XP86_SMIDL	Q9xp86 sminthopsis	337	41	52.6	567	2	Q59H79_HUMAN	Q59h79 homo sapien
265	42	53.8	381	2	O63538_9META	O63538 antechinus	338	41	52.6	619	2	Q9HTQ5_PSEAE	Q9htq5 pseudomonas
266	42	53.8	381	2	O63532_ANTSW	O63532 antechinus	339	41	52.6	658	2	Q4Q2C6_LEIMA	Q4q2c6 leishmania
267	42	53.8	381	2	Q9XP84_SMIGI	Q9xp84 sminthopsis	340	41	52.6	735	2	Q7YS57_BOVIN	Q7ys57 bos taurus
268	42	53.8	381	2	Q9XP87_SMIBI	Q9xp87 sminthopsis	341	41	52.6	763	2	Q7YS58_BOVIN	Q7ys58 bos taurus
269	42	53.8	381	2	Q9XP82_SMIGR	Q9xp82 sminthopsis	342	41	52.6	785	2	Q7YS56_BOVIN	Q7ys56 bos taurus
270	42	53.8	381	2	Q9XP81_SMIMI	Q9xp81 sminthopsis	343	41	52.6	796	2	Q541E5_MOUSE	Q541e5 mus musculu
271	42	53.8	381	2	Q35545_9META	Q35545 planigale s	344	41	52.6	826	2	Q7YS59_BOVIN	Q7ys59 bos taurus
272	42	53.8	381	2	Q9XP76_SMIPS	Q9xp76 sminthopsis	345	41	52.6	853	1	DNM3B_HUMAN	Q9ubc3 homo sapien
273	42	53.8	381	2	Q9XP85_SMIDO	Q9xp85 sminthopsis	346	41	52.6	859	1	DNM3B_MOUSE	O88509 mus musculu
274	42	53.8	381	2	Q9XP78_SMILO	Q9xp78 sminthopsis	347	41	52.6	869	2	Q6PI72_MOUSE	O6pi72 mus musculu
275	42	53.8	381	2	Q9XP79_SMILE	Q9xp79 sminthopsis	348	41	52.6	949	2	Q9ERT8_MOUSE	Q9ert8 mus musculu
276	42	53.8	381	2	O63533_ANTST	O63533 antechinus	349	41	52.6	949	2	Q7TQH5_MOUSE	Q7tqh5 mus musculu
277	42	53.8	381	2	Q33723_ANTHA	Q33723 antechinus	350	41	52.6	971	2	Q4QGF6_LEIMA	Q4qgf6 leishmania
278	42	53.8	381	2	Q9XP75_SMIVI	Q9xp75 sminthopsis	351	41	52.6	972	2	Q6ZWR6_MOUSE	Q6zwr6 mus musculu
279	42	53.8	381	2	Q5QS80_SMIDO	Q5qs80 sminthopsis	352	41	52.6	1025	2	Q9SAA5_ARATH	Q9saa5 arabidopsis
280	42	53.8	381	2	Q5QS54_PHATA	Q5qs54 phascogale	353	41	52.6	1117	2	Q8QP91_9HEPC	Q8qp91 hepatitis c
281	42	53.8	381	2	Q5J1T9_PHADO	Q5j1t9 phascolosor	354	41	52.6	1989	2	Q9ERT7_MOUSE	Q9ert7 mus musculu
282	42	53.8	381	2	Q5J1T8_PHADO	Q5j1t8 phascolosor	355	41	52.6	2225	1	PYR1_MESEAU	P08955 mesocricetu
283	42	53.8	399	2	Q38192_BPLC2	Q38192 lactococcus	356	41	52.6	2359	2	Q5YG24_XENTR	Q5yg24 xenopus tro
284	42	53.8	415	2	Q38311_EPLC2	Q38311 lactococcus	357	41	52.6	3010	2	Q6GYR9_9HEPC	Q6gyr9 hepatitis c
285	42	53.8	415	2	Q38445_9VIRU	Q38445 bacterioph	358	41	52.6	3010	2	Q5R2C0_9HEPC	Q5r2c0 hepatitis c
286	42	53.8	416	2	Q53CH4_9CAUD	Q53ch4 lactococcus	359	41	52.6	4342	2	Q9I157_PSEAE	Q9i157 pseudomonas
287	42	53.8	416	2	Q53CH8_9CAUD	Q53ch8 lactococcus	360	40.5	51.9	2331	2	Q59EG0_HUMAN	Q59eg0 homo sapien
288	42	53.8	416	2	Q53CI2_BPLC2	Q53ci2 lactococcus	361	40.5	51.9	4391	1	PGBM_HUMAN	P98160 homo sapien
289	42	53.8	692	2	Q55TZ6_CRYNE	Q55tz6 cryptococcu	362	40.5	51.9	4391	2	Q5VU27_HUMAN	Q5vu27 homo sapien
290	42	53.8	692	2	Q5KIE3_CRYNE	Q5kie3 cryptococcu	363	40	51.3	38	2	Q7YGP4_9HYST	Q7ygp4 ctenomys so
291	42	53.8	966	2	Q5GAI9_9VIRU	Q5gai9 grouper iri	364	40	51.3	59	2	Q85M86_9CETA	Q85m86 phocoena ph
292	42	53.8	968	2	Q5YFL3_9VIRU	Q5yfl3 singapore g	365	40	51.3	61	2	Q85M90_PHOPH	Q85m90 phocoena ph
293	42	53.8	1179	2	Q53K95_ORYSA	Q53k95 oryza sativ	366	40	51.3	63	2	Q85M91_9CETA	Q85m91 phocoenoide
294	42	53.8	1200	2	Q9VFH4_DROME	Q9vfh4 drosophila	367	40	51.3	66	2	Q85M92_9CETA	Q85m92 phocoenoide
295	41.5	52.6	107	2	Q7X9R3_GOSBA	Q7x9r3 gossypium b	368	40	51.3	76	2	Q85M88_9CETA	Q85m88 phocoena ph
296	41	52.6	75	2	Q8JZL6_MOUSE	Q8jzl6 mus musculu	369	40	51.3	76	2	Q85M87_9CETA	Q85m87 phocoena ph
297	41	52.6	103	2	Q7K6V2_CAEEL	Q7k6v2 caenorhabdi	370	40	51.3	77	2	Q85M93_9CETA	Q85m93 phocoenoide
298	41	52.6	131	2	Q7U7T5_SYNFX	Q7u7t5 synechococc	371	40	51.3	77	2	Q85M89_9CETA	Q85m89 phocoena ph
299	41	52.6	154	2	Q82S29_NITEU	Q82s29 nitrosomona	372	40	51.3	81	2	Q4SQZ9_TETNG	Q4sqz9 tetraodon n
300	41	52.6	161	2	Q7VD06_PROMA	Q7vd06 prochloroco	373	40	51.3	103	2	Q5UNR6_MICMA	Q5unr6 microtus mo
301	41	52.6	174	2	Q64683_MESAU	Q64683 mesocricetu	374	40	51.3	103	2	Q5UNR2_MICMA	Q5unr2 microtus mo
302	41	52.6	182	2	Q94QI5_CHIBR	Q94qi5 chinchilla	375	40	51.3	103	2	Q5UNR1_MICMA	Q5unr1 microtus mo
303	41	52.6	182	2	Q9B594_9HYST	Q9b594 lagidium vi	376	40	51.3	103	2	Q5UNR8_MICMA	Q5unr8 microtus mo
304	41	52.6	192	2	Q94QH4_AULBO	Q94qh4 auliscomys	377	40	51.3	105	2	Q6EX82_9RODE	Q6ex82 oecomys au
305	41	52.6	205	2	Q27672_9NEOP	Q27672 korscheltel	378	40	51.3	110	2	Q6EX83_9RODE	Q6ex83 oecomys bic
306	41	52.6	205	2	Q27771_9NEOP	Q27771 sthenopis a	379	40	51.3	117	2	Q9A9W6_CAUCR	Q9a9w6 caulobacter
307	41	52.6	218	2	Q9Z871_CHLPN	Q9z871 chlamydia p	380	40	51.3	120	2	Q5C8N2_9ADEN	Q5c8n2 simian aden
308	41	52.6	248	2	Q9ADH0_STRCO	Q9adh0 streptomyce	381	40	51.3	120	2	Q9T615_RASKA	Q9t615 rastrellige
309	41	52.6	264	2	Q98AK4_RHILO	Q98ak4 rhizobium l	382	40	51.3	121	2	Q76EP7_AUXRO	Q76ep7 auxis roche
310	41	52.6	267	2	Q71F27_9RODE	Q71f27 oxymycteru	383	40	51.3	121	2	Q76EP6_AUXRO	Q76ep6 auxis roche
311	41	52.6	292	2	Q7M9P2_WOLSU	Q7m9p2 wolinnella s	384	40	51.3	121	2	Q76EN9_AUXRO	Q76en9 auxis roche
312	41	52.6	294	2	Q93NG1_ARTNI	Q93ng1 arthrobacte	385	40	51.3	121	2	Q76EN8_AUXRO	Q76en8 auxis roche
313	41	52.6	306	2	Q8TWE7_METKA	Q8twe7 methanopyru	386	40	51.3	121	2	Q76EN2_AUXRO	Q76en2 auxis roche
314	41	52.6	309	2	Q9SSR0_ARATH	Q9ssr0 arabidopsis	387	40	51.3	121	2	Q76EF7_9SCOM	Q76ef7 auxis thaza
315	41	52.6	321	2	Q4S830_TETNG	Q4s830 tetraodon n	388	40	51.3	121	2	Q76ED1_AUXRO	Q76ed1 auxis roche
316	41	52.6	327	2	Q8M086_PROMN	Q8m086 prorocentru	389	40	51.3	121	2	Q76ED1_AUXRO	Q76ed1 auxis roche
317	41	52.6	351	2	Q5DE92_SCHJA	Q5de92 schistosoma	390	40	51.3	121	2	Q60GK4_THUAA	Q60gk4 thunnus ala
318	41	52.6	356	2	Q8M087_PROMN	Q8m087 prorocentru	391	40	51.3	121	2	Q60GK1_9SCOM	Q60gk1 euthynnus a
319	41	52.6	371	2	Q63HZ4_BURPS	Q63hz4 burkholderi	392	40	51.3	121	2	Q75ZP0_AUXRO	Q75zp0 auxis roche
320	41	52.6	371	2	Q629U8_BURMA	Q629u8 burkholderi	393	40	51.3	121	2	Q76EP8_AUXRO	Q76ep8 auxis roche
321	41	52.6	379	2	Q9T6J8_LEPCA	Q9t6j8 lepus capen	394	40	51.3	121	2	Q76EG0_9SCOM	Q76eg0 auxis thaza
322	41	52.6	379	2	Q9T6J1_LEPMN	Q9t6j1 lepus mand	395	40	51.3	121	2	Q60GK3_THUTH	Q60gk3 thunnus thy
323	41	52.6	379	2	Q9T6I5_LEPSI	Q9t6i5 lepus sinen	396	40	51.3	121	2		

397 Q60GK0 KATPE 121 2 Q60GK0 katsuwonus
398 Q9T612_SCOAU 125 2 Q9T612 scomber aus
399 Q79663_SCOAU 127 2 Q79663 scomber aus
400 Q9T611_SCOAU 127 2 Q9T611 scomber aus
401 Q9G2N0_9HYME 128 2 Q9G2N0 calomyrmex
402 P92842_9HYME 128 2 P92842 lysiphlebus
403 Q7YA05_MICAG 128 2 Q7YA05 microtus ag
404 Q7YA04_MICAG 128 2 Q7YA04 microtus ag
405 Q7Y7D3_MICAG 128 2 Q7Y7D3 microtus ag
406 Q7Y7L2_MICAG 128 2 Q7Y7L2 microtus ag
407 Q9MGJ7_MICLO 133 2 Q9MGJ7 microtus lo
408 Q9MGJ6_MICLO 136 2 Q9MGJ6 microtus lo
409 Q9MGJ8_MICLO 136 2 Q9MGJ8 microtus lo
410 Q9MGJ9_MICLO 136 2 Q9MGJ9 microtus lo
411 Q9ME21_MICLO 136 2 Q9ME21 microtus lo
412 Q9MDJ6_MICLO 136 2 Q9MDJ6 microtus lo
413 Q9T614_SCOSC 138 2 Q9T614 scomber sco
414 YMF28_MARPO 139 1 YMF28 marchantia
415 Q79661_RASKA 139 2 Q79661 rastrellige
416 Q85C83_9CETA 140 2 Q85C83 platanista
417 Q85PJ1_MOUSE 140 2 Q85PJ1 mus musculus
418 Q85PJ0_MESAU 140 2 Q85PJ0 mesocricetu
419 Q68GV8_9HYST 140 2 Q68GV8 hystrix ind
420 Q9T613_SCOJP 140 2 Q9T613 scomber jap
421 Q79664_SCOAU 140 2 Q79664 scomber aus
422 Q79662_SCOJP 140 2 Q79662 scomber jap
423 Q6PVU8_9HYME 144 2 Q6PVU8 trigona hoc
424 Q4VVK9_9PERC 151 2 Q4VVK9 iciphthys l
425 Q9T4Z6_PHYDA 155 2 Q9T4Z6 phyllotis d
426 Q9G987_VESCR 163 2 Q9G987 vespa crabr
427 Q94QI1_9HYST 169 2 Q94QI1 octomys mim
428 Q8HIW2_BARBR 170 2 Q8HIW2 barbastella
429 Q8HIH3_PLEAS 170 2 Q8HIH3 plecotus au
430 Q8HIH1_9CHIR 170 2 Q8HIH1 plecotus te
431 Q8HIH0_9CHIR 170 2 Q8HIH0 plecotus te
432 Q8HIG9_9CHIR 170 2 Q8HIG9 plecotus te
433 Q8HBJ5_9CHIR 170 2 Q8HBJ5 plecotus te
434 Q8HB72_BARBR 170 2 Q8HB72 barbastella
435 Q9T2F4_LAGAC 173 2 Q9T2F4 lagenorhync
436 Q9T2F5_LAGAL 173 2 Q9T2F5 lagenorhync
437 Q85ID5_9SAUR 173 2 Q85ID5 liolaemus p
438 Q85ID4_9SAUR 173 2 Q85ID4 liolaemus e
439 Q85ID2_9SAUR 173 2 Q85ID2 liolaemus p
440 Q85IC9_9SAUR 173 2 Q85IC9 liolaemus v
441 Q85IC8_9SAUR 173 2 Q85IC8 liolaemus s
442 Q85IC7_9SAUR 173 2 Q85IC7 liolaemus s
443 Q85IC6_9SAUR 173 2 Q85IC6 liolaemus s
444 Q85IC5_9SAUR 173 2 Q85IC5 liolaemus s
445 Q85IC3_9SAUR 173 2 Q85IC3 liolaemus k
446 Q85IC1_9SAUR 173 2 Q85IC1 liolaemus p
447 Q85IB9_9SAUR 173 2 Q85IB9 liolaemus p
448 Q85IB8_9SAUR 173 2 Q85IB8 liolaemus e
449 Q85IB7_9SAUR 173 2 Q85IB7 liolaemus l
450 Q85IB6_9SAUR 173 2 Q85IB6 liolaemus s
451 Q85ID1_9SAUR 173 2 Q85ID1 liolaemus s
452 Q85CA5_9SAUR 173 2 Q85CA5 liolaemus s
453 Q85C90_9SAUR 173 2 Q85C90 liolaemus s
454 Q85B28_9SAUR 173 2 Q85B28 liolaemus p
455 Q85AP0_9SAUR 173 2 Q85AP0 liolaemus e
456 Q85AN9_9SAUR 173 2 Q85AN9 liolaemus b
457 Q85AN8_9SAUR 173 2 Q85AN8 liolaemus e
458 Q85AN7_9SAUR 173 2 Q85AN7 liolaemus e
459 Q85AN6_9SAUR 173 2 Q85AN6 liolaemus k
460 Q85AN5_9SAUR 173 2 Q85AN5 liolaemus s
461 Q85AN4_9SAUR 173 2 Q85AN4 liolaemus s
462 Q85AN3_9SAUR 173 2 Q85AN3 liolaemus s
463 Q85AN2_9SAUR 173 2 Q85AN2 liolaemus s
464 Q85AN1_9SAUR 173 2 Q85AN1 liolaemus s
465 Q85ANO_9SAUR 173 2 Q85ANO liolaemus k
466 Q85CL4_9SAUR 173 2 Q85CL4 liolaemus p
467 Q85CI9_9SAUR 173 2 Q85CI9 liolaemus e
468 Q85CI8_9SAUR 173 2 Q85CI8 liolaemus s
469 Q85CI7_9SAUR 173 2 Q85CI7 liolaemus s

470 Q85AK8_9SAUR 173 2 Q85AK8_9SAUR
471 Q85AK7_9SAUR 173 2 Q85AK7_9SAUR
472 Q85CH1_9SAUR 173 2 Q85CH1_9SAUR
473 Q85CH0_9SAUR 173 2 Q85CH0_9SAUR
474 Q85AY8_9SAUR 173 2 Q85AY8_9SAUR
475 Q85CB4_9SAUR 173 2 Q85CB4_9SAUR
476 Q85AD8_9SAUR 173 2 Q85AD8_9SAUR
477 Q85DJ4_APOSY 174 2 Q85DJ4_APOSY
478 Q85DJ0_APOSY 174 2 Q85DJ0_APOSY
479 CYB_IDIPH 176 1 CYB_IDIPH
480 CYB_SCICA 176 1 CYB_SCICA
481 Q656A5_ORYSA 176 2 Q656A5_ORYSA
482 Q36131_TAMST 176 2 Q36131_TAMST
483 Q35955_9SCIU 176 2 Q35955_9SCIU
484 Q34105_CYNGU 176 2 Q34105_CYNGU
485 Q33725_9SCIU 176 2 Q33725_9SCIU
486 Q34983_9SCIU 176 2 Q34983_9SCIU
487 Q35913_SPESP 176 2 Q35913_SPESP
488 Q34495_TAMQU 176 2 Q34495_TAMQU
489 Q35931_9SCIU 176 2 Q35931_9SCIU
490 Q35824_SPEFR 176 2 Q35824_SPEFR
491 Q33714_AMMHA 176 2 Q33714_AMMHA
492 Q34482_TAMMI 176 2 Q34482_TAMMI
493 Q35894_SCINI 176 2 Q35894_SCINI
494 Q34415_TAMCA 176 2 Q34415_TAMCA
495 Q34423_TAMDO 176 2 Q34423_TAMDO
496 Q34131_CYNLU 176 2 Q34131_CYNLU
497 Q7GG55_GLASA 176 2 Q7GG55_GLASA
498 Q85DJ2_APOSY 177 2 Q85DJ2_APOSY
499 Q9B595_CHILA 178 2 Q9B595_CHILA
500 Q4VWI3_9RODE 178 2 Q4VWI3_9RODE
501 Q94QI6_CHIBR 179 2 Q94QI6_CHIBR
502 Q9T2F3_LAGOB 180 2 Q9T2F3_LAGOB
503 Q9T2F1_LAGOB 180 2 Q9T2F1_LAGOB
504 Q9T2F0_LAGOL 180 2 Q9T2F0_LAGOL
505 Q85DJ9_APOSY 180 2 Q85DJ9_APOSY
506 Q9XNN8_LAGAC 181 2 Q9XNN8_LAGAC
507 Q9XNN9_LAGAL 181 2 Q9XNN9_LAGAL
508 Q85GL0_9HYST 181 2 Q85GL0_9HYST
509 Q5MAZ9_9SAUR 181 2 Q5MAZ9_9SAUR
510 Q5MAY4_9SAUR 181 2 Q5MAY4_9SAUR
511 Q5MAY2_9SAUR 181 2 Q5MAY2_9SAUR
512 Q5MAY0_9SAUR 181 2 Q5MAY0_9SAUR
513 Q7YFF3_MARMU 182 2 Q7YFF3_MARMU
514 Q94QI3_CHILA 182 2 Q94QI3_CHILA
515 Q94QI4_CHILA 182 2 Q94QI4_CHILA
516 Q94QI2_CHILA 182 2 Q94QI2_CHILA
517 Q9B592_CHIBR 182 2 Q9B592_CHIBR
518 Q8HH86_CHILA 182 2 Q8HH86_CHILA
519 Q4VTW0_9SAUR 182 2 Q4VTW0_9SAUR
520 Q9T2F2_LAGOB 185 2 Q9T2F2_LAGOB
521 Q68K94_AMMLU 185 2 Q68K94_AMMLU
522 Q68K85_AMMLU 185 2 Q68K85_AMMLU
523 Q68K71_AMMLU 185 2 Q68K71_AMMLU
524 Q68KE1_AMMLU 185 2 Q68KE1_AMMLU
525 Q68KA9_AMMLU 185 2 Q68KA9_AMMLU
526 Q68K91_AMMLU 185 2 Q68K91_AMMLU
527 Q9ZZK5_9SAUR 185 2 Q9ZZK5_9SAUR
528 Q9T510_UTAST 185 2 Q9T510_UTAST
529 Q9T520_UTAST 185 2 Q9T520_UTAST
530 Q9ZZK9_9SAUR 185 2 Q9ZZK9_9SAUR
531 Q9T509_UTAST 185 2 Q9T509_UTAST
532 Q9T517_UTAST 185 2 Q9T517_UTAST
533 Q9ZZK7_9SAUR 185 2 Q9ZZK7_9SAUR
534 Q9T515_UTAST 185 2 Q9T515_UTAST
535 Q9ZZK3_UTAST 185 2 Q9ZZK3_UTAST
536 Q9T519_UTAST 185 2 Q9T519_UTAST
537 Q9T518_UTAST 185 2 Q9T518_UTAST
538 Q9T516_UTAST 185 2 Q9T516_UTAST
539 Q9T514_UTAST 185 2 Q9T514_UTAST
540 Q9T512_UTAST 185 2 Q9T512_UTAST
541 Q9ZZK8_9SAUR 185 2 Q9ZZK8_9SAUR
542

Q85ak8 liolaemus e
Q85ak7 liolaemus s
Q85ch1 liolaemus p
Q85ch0 liolaemus p
Q85ay8 liolaemus p
Q85cb4 liolaemus p
Q85ad8 liolaemus s
Q85dj4 apodemus sy
Q85dj0 apodemus sy
Q33537 idionycteri
Q35805 sciurus car
Q656a5 oryza sativ
Q36131 tamias stri
Q35955 spermophilu
Q34105 cynomys gun
Q33725 ammospermop
Q34983 microsciuru
Q35913 spermophilu
Q34495 tamias quad
Q35931 sciurus str
Q35824 spermophilu
Q33714 ammospermop
Q34482 tamias mini
Q35894 sciurus nig
Q34415 tamias cani
Q34423 tamias dors
Q34131 cynomys lud
Q7GG55 glaucomys s
Q85dj2 apodemus sy
Q9b595 chinchilla
Q4vwi3 thalpomys l
Q94qi6 chinchilla
Q9t2f3 lagenorhync
Q9t2f1 lagenorhync
Q9t2f0 lagenorhync
Q85dj9 apodemus sy
Q9xnn8 lagenorhync
Q9xnn9 lagenorhync
Q85gl0 cavia tschu
Q5maz9 tarentola b
Q5may4 tarentola b
Q5may2 tarentola b
Q5may0 tarentola b
Q5may1 tarentola b
Q7yff3 marmosa mur
Q94qi3 chinchilla
Q94qi4 chinchilla
Q94qi2 chinchilla
Q9b592 chinchilla
Q8bh86 chinchilla
Q4vtw0 gymnodactyl
Q9t2f2 lagenorhync
Q68k94 ammospermop
Q68k85 ammospermop
Q68k71 ammospermop
Q68ke1 ammospermop
Q68ka9 ammospermop
Q68k91 ammospermop
Q9zzk5 uta squamat
Q9t510 uta stansbu
Q9t520 uta stansbu
Q9zzk9 urosaurus m
Q9t509 uta stansbu
Q9t517 uta stansbu
Q9zzk7 uta palmeri
Q9t515 uta stansbu
Q9zzk3 uta stansbu
Q9t519 uta stansbu
Q9t518 uta stansbu
Q9t516 uta stansbu
Q9t514 uta stansbu
Q9t512 uta stansbu
Q9zzk8 petrosaurus

543 40 51.3 185 2 Q9T508 UTAST Q9t508 uta stansbu 616 40 51.3 195 2 Q85QY5_LEPCA 1045QY5 lepus capen
544 40 51.3 185 2 Q9ZZK2_9SAUR Q9zzk2 uta stejnæg 617 40 51.3 195 2 Q85QY3_LEPCA 1045QY3 lepus capen
545 40 51.3 185 2 Q9T513_UTAST Q9t513 uta stansbu 618 40 51.3 195 2 Q85QY2_LEPCA 1045QY2 lepus capen
546 40 51.3 185 2 Q9ZZK6_9SAUR Q9zzk6 uta stellat 619 40 51.3 195 2 Q85QY1_LEPSX 1045QY1 lepus saxat
547 40 51.3 185 2 Q9ZZK4_9SAUR Q9zzk4 uta antiqua 620 40 51.3 195 2 Q85QY0_LEPSX 1045QY0 lepus saxat
548 40 51.3 185 2 Q9T511_UTAST Q9t511 uta stansbu 621 40 51.3 195 2 Q85BA6_LEPGR 1045BA6 lepus grana
549 40 51.3 185 2 Q9T3A3_UTAST Q9t3a3 uta stansbu 622 40 51.3 195 2 Q85B32_LEPGR 1045B32 lepus grana
550 40 51.3 185 2 Q9T3A2_UTAST Q9t3a2 uta stansbu 623 40 51.3 195 2 Q85B31_LEPEU 1045B31 lepus europa
551 40 51.3 185 2 Q4VTV4_9SAUR Q4vtv4 gymnodactyl 624 40 51.3 195 2 Q36161_THUTH 1036161 thunnus thy
552 40 51.3 186 2 Q9XNN7_LAGOL Q9xnn7 lagenorhync 625 40 51.3 196 2 Q33879_9SCOM 1033879 auxis thaza
553 40 51.3 186 2 Q85Q29_LEPEU Q85q29 lepus europa 626 40 51.3 196 2 Q34405_9SCOM 1034405 euthynnus a
554 40 51.3 186 2 Q5VC40_9RODE Q5vc40 peromyscus 627 40 51.3 196 2 Q34403_9SCOM 1034403 euthynnus a
555 40 51.3 186 2 Q5VBUI_9RODE Q5vbu1 peromyscus 628 40 51.3 197 2 Q4VTX0_9SAUR 104VTX0_9SAUR
556 40 51.3 186 2 Q5VBS9_9RODE Q5vbs9 peromyscus 629 40 51.3 198 2 Q85LL0_9SCOM 1085LL0_9SCOM
557 40 51.3 186 2 Q5VBQ1_9RODE Q5vbk1 peromyscus 630 40 51.3 198 2 Q85LK7_9SCOM 1085LK7_9SCOM
558 40 51.3 186 2 Q8HBP8_CHIBR Q8hbp8 chinchilla 631 40 51.3 198 2 Q85LK5_KATPE 1085LK5_katsuwonu
559 40 51.3 186 2 Q5VCA6_9RODE Q5vca6 peromyscus 632 40 51.3 198 2 Q85LK3_SAROR 1085LK3_sarda orien
560 40 51.3 186 2 Q5VBM2_9RODE Q5vbm2 peromyscus 633 40 51.3 198 2 Q85LK1_THUAA 1085LK1_thunnus ala
561 40 51.3 186 2 Q5VBP8_9RODE Q5vbp8 peromyscus 634 40 51.3 198 2 Q85LJ8_THUAL 1085LJ8_thunnus alb
562 40 51.3 186 2 Q5VBS6_9RODE Q5vbs6 peromyscus 635 40 51.3 198 2 Q85CG0_AUXRO 1085CG0_auxis roche
563 40 51.3 186 2 Q5VBL7_9RODE Q5vbl7 peromyscus 636 40 51.3 198 2 Q85CF9_KATPE 1085CF9_katsuwonu
564 40 51.3 186 2 Q5VBR3_9RODE Q5vbr3 peromyscus 637 40 51.3 198 2 Q4VTX6_9SAUR 104VTX6_gymnodactyl
565 40 51.3 186 2 Q5VBT9_9RODE Q5vbt9 peromyscus 638 40 51.3 198 2 Q4VTV6_9SAUR 104VTV6_gymnodactyl
566 40 51.3 186 2 Q5VBL0_9RODE Q5vbl0 peromyscus 639 40 51.3 199 1 CVB_SARCH 10CVB_SARCH
567 40 51.3 186 2 Q5VBM9_9RODE Q5vbm9 peromyscus 640 40 51.3 199 2 Q7YFF2_MARMU 107YFF2_marmosa mur
568 40 51.3 186 2 Q5VBS7_9RODE Q5vbs7 peromyscus 641 40 51.3 199 2 Q35975_THUAA 1035975_thunnus ala
569 40 51.3 187 2 Q85QY7_LEPTI Q85qy7 lepus timid 642 40 51.3 199 2 Q35867_SCOJP 1035867_scomber jap
570 40 51.3 187 2 Q6TQW3_9SAUR Q6tqw3 liolaemus g 643 40 51.3 199 2 Q34836_KATPE 1034836_katsuwonu
571 40 51.3 187 2 Q4VTY4_9SAUR Q4vtY4 gymnodactyl 644 40 51.3 199 2 Q34590_9SCOM 1034590_gasterochis
572 40 51.3 187 2 Q4VTX9_9SAUR Q4vtx9 gymnodactyl 645 40 51.3 200 2 Q9ZZG3_9RODE 109ZZG3_phyllotis m
573 40 51.3 190 2 Q85QY6_LEPTI Q85qy6 lepus timid 646 40 51.3 200 2 Q85A48_CLEGL 1085A48_clethrionom
574 40 51.3 190 2 Q9MGJ4_9SAUR Q9mgj4 darevskia v 647 40 51.3 200 2 Q52TK6_CLEGA 1052TK6_clethrionom
575 40 51.3 190 2 Q4VTV1_9SAUR Q4vtv1 gymnodactyl 648 40 51.3 200 2 Q52TJ4_CLEGA 1052TJ4_clethrionom
576 40 51.3 191 2 Q85QZ8_LEPEU Q85qz8 lepus europa 649 40 51.3 200 2 Q52TJ2_CLEGA 1052TJ2_clethrionom
577 40 51.3 191 2 Q85B70_LEPGR Q85b70 lepus grana 650 40 51.3 200 2 Q52TJ0_CLEGA 1052TJ0_clethrionom
578 40 51.3 192 2 Q33689_9CETA Q33689 australopho 651 40 51.3 200 2 Q52TI9_CLEGA 1052TI9_clethrionom
579 40 51.3 192 2 Q35164_NEOPH Q35164 neophocaena 652 40 51.3 200 2 Q52TI8_CLEGA 1052TI8_clethrionom
580 40 51.3 192 2 Q35628_PHOSS Q35628 phocoena si 653 40 51.3 200 2 Q52TI3_CLEGA 1052TI3_clethrionom
581 40 51.3 192 2 Q35603_PROPH Q35603 phocoenoide 654 40 51.3 200 2 Q52TH5_CLEGA 1052TH5_clethrionom
582 40 51.3 192 2 Q85R09_LEPGR Q85r09 lepus grana 655 40 51.3 200 2 Q52TJ5_CLEGA 1052TJ5_clethrionom
583 40 51.3 192 2 Q85QZ1_LEPCR Q85qz1 lepus cori 656 40 51.3 200 2 Q52TI7_CLEGA 1052TI7_clethrionom
584 40 51.3 192 2 Q37055_9CETA Q37055 phocoenoide 657 40 51.3 200 2 Q52TH6_CLEGA 1052TH6_clethrionom
585 40 51.3 192 2 Q9MPW6_CHILA Q9mpw6 chinchilla 658 40 51.3 200 2 Q52TM9_CLEGA 1052TM9_clethrionom
586 40 51.3 192 2 Q6X4W7_9HYST Q6x4w7 lagidium vi 659 40 51.3 200 2 Q52TK9_CLEGA 1052TK9_clethrionom
587 40 51.3 193 2 Q85R10_LEPGR Q85r10 lepus grana 660 40 51.3 200 2 Q52TK9_CLEGA 1052TK9_clethrionom
588 40 51.3 193 2 Q85QZ7_LEPEU Q85qz7 lepus europa 661 40 51.3 200 2 Q52TK9_CLEGA 1052TK9_clethrionom
589 40 51.3 193 2 Q85QZ4_LEPEU Q85qz4 lepus europa 662 40 51.3 201 2 Q9G4Q6_9RODE 109G4Q6_peromyscus
590 40 51.3 193 2 Q85QZ2_LEPCS Q85qz2 lepus castr 663 40 51.3 201 2 Q33861_9SCOM 1033861_acanthocybi
591 40 51.3 193 2 Q85QY9_LEPTI Q85qy9 lepus timid 664 40 51.3 201 2 Q85I76_9SAUR 1085I76_liolaemus s
592 40 51.3 193 2 Q6X4W6_9HYST Q6x4w6 lagidium vi 665 40 51.3 202 2 Q36093_THUOB 1036093_thunnus obe
593 40 51.3 194 2 Q85R13_LEPGR Q85r13 lepus grana 666 40 51.3 202 2 Q6TQV6_9SAUR 106TQV6_liolaemus g
594 40 51.3 194 2 Q85R05_LEPGR Q85r05 lepus grana 667 40 51.3 203 2 Q35923_SCOSC 1035923_scomber sco
595 40 51.3 194 2 Q85R04_LEPGR Q85r04 lepus grana 668 40 51.3 203 2 Q85I99_9SAUR 1085I99_liolaemus e
596 40 51.3 194 2 Q85R03_LEPGR Q85r03 lepus grana 669 40 51.3 204 2 Q27639_9NEOP 1027639_epimartyria
597 40 51.3 194 2 Q5IXI6_9SAUR Q5ixi6 thecadactyl 670 40 51.3 204 2 Q4VTY3_9SAUR 104VTY3_gymnodactyl
598 40 51.3 194 2 Q5IXG9_9SAUR Q5ixg9 thecadactyl 671 40 51.3 205 2 Q9MG44_9SCOM 109MG44_thunnus atl
599 40 51.3 194 2 Q5IXG8_9SAUR Q5ixg8 thecadactyl 672 40 51.3 205 2 Q85LL1_AUXRO 1085LL1_auxis roche
600 40 51.3 194 2 Q5IXG2_9SAUR Q5ixg2 thecadactyl 673 40 51.3 205 2 Q85LK9_9SCOM 1085LK9_auxis thaza
601 40 51.3 194 2 Q5IXG1_9SAUR Q5ixg1 thecadactyl 674 40 51.3 205 2 Q85LK6_9SCOM 1085LK6_euthynnus a
602 40 51.3 194 2 Q4VTW2_9SAUR Q4vtw2 gymnodactyl 675 40 51.3 205 2 Q85LK4_KATPE 1085LK4_katsuwonu
603 40 51.3 195 2 Q85R14_LEPGR Q85r14 lepus grana 676 40 51.3 205 2 Q85LK2_SAROR 1085LK2_sarda orien
604 40 51.3 195 2 Q85R12_LEPGR Q85r12 lepus grana 677 40 51.3 205 2 Q85LJ7_THUAL 1085LJ7_thunnus alb
605 40 51.3 195 2 Q85R11_LEPGR Q85r11 lepus grana 678 40 51.3 205 2 Q85LJ4_9SCOM 1085LJ4_thunnus mac
606 40 51.3 195 2 Q85R07_LEPGR Q85r07 lepus grana 679 40 51.3 205 2 Q85LJ3_THUOB 1085LJ3_thunnus obe
607 40 51.3 195 2 Q85R06_LEPGR Q85r06 lepus grana 680 40 51.3 205 2 Q85LJ0_THUOB 1085LJ0_thunnus obe
608 40 51.3 195 2 Q85R02_LEPGR Q85r02 lepus grana 681 40 51.3 205 2 Q85LI8_THUTO 1085LI8_thunnus thy
609 40 51.3 195 2 Q85R01_LEPGR Q85r01 lepus grana 682 40 51.3 205 2 Q85LI6_THUTH 1085LI6_thunnus thy
610 40 51.3 195 2 Q85R00_LEPEU Q85r00 lepus europa 683 40 51.3 205 2 Q9MG43_9SCOM 109MG43_thunnus ton
611 40 51.3 195 2 Q85QZ6_LEPEU Q85qz6 lepus europa 684 40 51.3 206 2 Q36069_9SCOM 1036069_thunnus mac
612 40 51.3 195 2 Q85QZ5_LEPEU Q85qz5 lepus europa 685 40 51.3 206 2 Q6TQY1_9SAUR 106TQY1_liolaemus s
613 40 51.3 195 2 Q85QZ3_LEPCS Q85qz3 lepus castr 686 40 51.3 206 2 Q6TQV5_9SAUR 106TQV5_liolaemus g
614 40 51.3 195 2 Q85QZ0_LEPTI Q85qz0 lepus timid 687 40 51.3 207 2 Q6X4W9_9HYST 106X4W9_lagidium pe
615 40 51.3 195 2 Q85QY8_LEPTI Q85qy8 lepus timid 688 40 51.3 207 2 Q6X4W8_9HYST 106X4W8_lagidium vi

1045QY5 lepus capen
1045QY3 lepus capen
1045QY2 lepus capen
1045QY1 lepus saxat
1045QY0 lepus saxat
1045ba6 lepus grana
1045b32 lepus grana
1045b31 lepus europa
1036161 thunnus thy
1033879 auxis thaza
1034405 euthynnus a
1034403 euthynnus a
104vtx0 gymnodactyl
104vtx0 auxis thaza
1085lk7 euthynnus a
1085lk5 katsuwonu
1085lk3 sarda orien
1085lk1 thunnus ala
1085lj8 thunnus alb
1085cg0 auxis roche
1085cf9 katsuwonu
104vtx6 gymnodactyl
104vtv6 gymnodactyl
1034864 sarda chili
107yff2 marmosa mur
1035975 thunnus ala
1035867 scomber jap
1034836 katsuwonu
1034590 gasterochis
109zzg3 phyllotis m
1085a48 clethrionom
1052tk6 clethrionom
1052tj4 clethrionom
1052tj2 clethrionom
1052tj0 clethrionom
1052ti9 clethrionom
1052ti8 clethrionom
1052ti3 clethrionom
1052th5 clethrionom
1052tj5 clethrionom
1052ti7 clethrionom
1052th6 clethrionom
1052tm9 clethrionom
1052tk9 clethrionom
1035976 thunnus alb
1085ia4 liolaemus e
109g4q6 peromyscus
1033861 acanthocybi
1085i76 liolaemus s
1036093 thunnus obe
106tqv6 liolaemus g
1035923 scomber sco
1085i99 liolaemus e
1027639 epimartyria
104vtY3 gymnodactyl
109mg44 thunnus atl
1085ll1 auxis roche
1085lk9 auxis thaza
1085lk6 euthynnus a
1085lk4 katsuwonu
1085lk2 sarda orien
1085lj7 thunnus alb
1085lj4 thunnus mac
1085lj3 thunnus obe
1085lj0 thunnus obe
1085li8 thunnus thy
1085li6 thunnus thy
109mg43 thunnus ton
1036069 thunnus mac
106tqy1 liolaemus s
106tqv5 liolaemus g
106x4w9 lagidium pe
106x4w8 lagidium vi

689 Q6X4W5 lagidium vi 207 2 Q6X4W5_9HYST 40 51.3

690 Q85IB0 liolaemus e 207 2 Q85IB0_9SAUR 40 51.3

691 Q85IA7 liolaemus e 207 2 Q85IA7_9SAUR 40 51.3

692 Q6TR01 liolaemus d 207 2 Q6TR01_9SAUR 40 51.3

693 Q21639 darevskia d 208 2 Q21639_9SAUR 40 51.3

694 Q21632 darevskia a 208 2 Q21632_9SAUR 40 51.3

695 Q58J79 lepus grana 209 2 Q58J79_LEPGR 40 51.3

696 Q58J78 lepus grana 209 2 Q58J78_LEPEU 40 51.3

697 Q58J77 lepus europ 209 2 Q58J77_LEPEU 40 51.3

698 Q58J76 lepus grana 209 2 Q58J76_LEPGR 40 51.3

699 Q58J75 lepus castr 209 2 Q58J75_LEPCS 40 51.3

700 Q52N14 mabuya macu 209 2 Q52N14_9SAUR 40 51.3

701 Q4VTW6 gymnodactyl 210 2 Q4VTW6_9SAUR 40 51.3

702 Q85LK0 thunnus ala 211 2 Q85LK0_THUAA 40 51.3

703 Q85LJ9 thunnus alb 211 2 Q85LJ9_THUAL 40 51.3

704 Q85LJ6 thunnus atl 211 2 Q85LJ6_9SCOM 40 51.3

705 Q85LJ5 thunnus mac 211 2 Q85LJ5_9SCOM 40 51.3

706 Q85LJ2 thunnus obe 211 2 Q85LJ2_THUOB 40 51.3

707 Q85LJ1 thunnus obe 211 2 Q85LJ1_THUOB 40 51.3

708 Q85LI9 thunnus thy 211 2 Q85LI9_THUTO 40 51.3

709 Q85LI7 thunnus thy 211 2 Q85LI7_THUTH 40 51.3

710 Q85LI5 thunnus ton 211 2 Q85LI5_9SCOM 40 51.3

711 Q85LL2 auxis roche 212 2 Q85LL2_AUXRO 40 51.3

712 Q6TR10 liolaemus d 212 2 Q6TR10_9SAUR 40 51.3

713 Q52N09 mabuya macu 212 2 Q52N09_9SAUR 40 51.3

714 Q69FV4 arborimus l 213 2 Q69FV4_9RODE 40 51.3

715 Q9MGJ3 darevskia p 213 2 Q9MGJ3_9SAUR 40 51.3

716 Q7YGM2 kenopia str 213 2 Q7YGM2_9PASS 40 51.3

717 Q6TR40 liolaemus d 213 2 Q6TR40_9SAUR 40 51.3

718 Q6TQY0 liolaemus s 213 2 Q6TQY0_9SAUR 40 51.3

719 Q6TQV3 liolaemus g 213 2 Q6TQV3_9SAUR 40 51.3

720 Q52N10 mabuya macu 213 2 Q52N10_9SAUR 40 51.3

721 Q85TD9 cavia tschu 214 2 Q85TD9_9HYST 40 51.3

722 Q34872 oryctolagus 215 2 Q34872_RABIT 40 51.3

723 Q85TD3 galea muste 215 2 Q85TD3_GALMU 40 51.3

724 Q52N11 mabuya macu 215 2 Q52N11_9SAUR 40 51.3

725 Q52N03 mabuya macu 215 2 Q52N03_9SAUR 40 51.3

726 Q9TAG3 uma paraphy 216 2 Q9TAG3_9SAUR 40 51.3

727 Q9TAG4 uma exsul. 216 2 Q9TAG4_9SAUR 40 51.3

728 Q4VTY0 gymnodactyl 217 2 Q4VTY0_9SAUR 40 51.3

729 Q9TAF7 holbrookia 217 2 Q9TAF7_HOLMA 40 51.3

730 Q9TAF3 holbrookia 217 2 Q9TAF3_9SAUR 40 51.3

731 Q9TAG7 copte 217 2 Q9TAG7_COPT 40 51.3

732 Q9TAG6 copte 217 2 Q9TAG6_COPT 40 51.3

733 Q9TAG2 9SAUR 217 2 Q9TAG2_9SAUR 40 51.3

734 Q9TAF4 holma 217 2 Q9TAF4_HOLMA 40 51.3

735 Q9TAG0 holma 217 2 Q9TAG0_HOLMA 40 51.3

736 Q9TAH3 uroor 217 2 Q9TAH3_UROOR 40 51.3

737 Q9TAH4 uta stansbu 217 2 Q9TAH4_UTAST 40 51.3

738 Q9TAF6 holbrookia 217 2 Q9TAF6_HOLMA 40 51.3

739 Q9TAG8 calldr 217 2 Q9TAG8_CALDR 40 51.3

740 Q9TAF8 holma 217 2 Q9TAF8_HOLMA 40 51.3

741 Q9TAG9 calldr 217 2 Q9TAG9_CALDR 40 51.3

742 Q9TAF9 holma 217 2 Q9TAF9_HOLMA 40 51.3

743 Q9TAF5 holbrookia 217 2 Q9TAF5_HOLMA 40 51.3

744 Q9TAG1 9SAUR 217 2 Q9TAG1_9SAUR 40 51.3

745 Q9TAG5 9SAUR 217 2 Q9TAG5_9SAUR 40 51.3

746 Q6TQV0 liolaemus o 217 2 Q6TQV0_9SAUR 40 51.3

747 Q9T454 uma scopari 217 2 Q9T454_UMASC 40 51.3

748 Q9T4Q0 callisaur 217 2 Q9T4Q0_CALDR 40 51.3

749 Q9T3Y7 uma notata 217 2 Q9T3Y7_UMANO 40 51.3

750 Q6TR09 liolaemus d 218 2 Q6TR09_9SAUR 40 51.3

751 Q6TQV4 liolaemus g 218 2 Q6TQV4_9SAUR 40 51.3

752 Q4VTV8 gymnodactyl 218 2 Q4VTV8_9SAUR 40 51.3

753 Q6KCT8 micoureus d 219 2 Q6KCT8_MICDE 40 51.3

754 Q9MM61 eremias vel 219 2 Q9MM61_9SAUR 40 51.3

755 Q6TQX8 liolaemus s 219 2 Q6TQX8_9SAUR 40 51.3

756 Q6RBY2 marmota mon 220 2 Q6RBY2_MARMO 40 51.3

757 Q9MIB1 mabuya quin 220 2 Q9MIB1_9SAUR 40 51.3

758 Q85NU8 pachydactyl 220 2 Q85NU8_9SAUR 40 51.3

759 Q85H40 pachydactyl 220 2 Q85H40_9SAUR 40 51.3

760 Q6TR56 liolaemus d 220 2 Q6TR56_9SAUR 40 51.3

761 Q4VTV5 gymnodactyl 220 2 Q4VTV5_9SAUR 40 51.3

762 Q6X4W5 40 51.3

763 Q85IB0 liolaemus e 40 51.3

764 Q85IA7 liolaemus e 40 51.3

765 Q6TR01 liolaemus d 40 51.3

766 Q21639 darevskia d 40 51.3

767 Q21632 darevskia a 40 51.3

768 Q58J79 lepus grana 40 51.3

769 Q58J78 lepus grana 40 51.3

770 Q58J77 lepus europ 40 51.3

771 Q58J76 lepus grana 40 51.3

772 Q58J75 lepus castr 40 51.3

773 Q52N14 mabuya macu 40 51.3

774 Q4VTW6 gymnodactyl 40 51.3

775 Q85LK0 thunnus ala 40 51.3

776 Q85LJ9 thunnus alb 40 51.3

777 Q85LJ6 thunnus atl 40 51.3

778 Q85LJ5 thunnus mac 40 51.3

779 Q85LJ2 thunnus obe 40 51.3

780 Q85LJ1 thunnus obe 40 51.3

781 Q85LI9 thunnus thy 40 51.3

782 Q85LI7 thunnus thy 40 51.3

783 Q85LI5 thunnus ton 40 51.3

784 Q85LL2 auxis roche 40 51.3

785 Q6TR10 liolaemus d 40 51.3

786 Q52N09 mabuya macu 40 51.3

787 Q69FV4 arborimus l 40 51.3

788 Q9MGJ3 darevskia p 40 51.3

789 Q7YGM2 kenopia str 40 51.3

790 Q6TR40 liolaemus d 40 51.3

791 Q6TQY0 liolaemus s 40 51.3

792 Q6TQV3 liolaemus g 40 51.3

793 Q52N10 mabuya macu 40 51.3

794 Q85TD9 cavia tschu 40 51.3

795 Q34872 oryctolagus 40 51.3

796 Q85TD3 galea muste 40 51.3

797 Q52N11 mabuya macu 40 51.3

798 Q52N03 mabuya macu 40 51.3

799 Q9TAG3 uma paraphy 40 51.3

800 Q9TAG4 uma exsul. 40 51.3

801 Q4VTY0 gymnodactyl 40 51.3

802 Q9TAF7 holbrookia 40 51.3

803 Q9TAF3 holbrookia 40 51.3

804 Q9TAG7 coptosaurus 40 51.3

805 Q9TAG6 coptosaurus 40 51.3

806 Q9TAG2 holbrookia 40 51.3

807 Q9TAF4 holbrookia 40 51.3

808 Q9TAG0 holbrookia 40 51.3

809 Q9TAH3 urosaurus o 40 51.3

810 Q9TAH4 uta stansbu 40 51.3

811 Q9TAF6 holbrookia 40 51.3

812 Q9TAG8 callisaur 40 51.3

813 Q9TAF8 holbrookia 40 51.3

814 Q9TAG9 callisaur 40 51.3

815 Q9TAF9 holbrookia 40 51.3

816 Q9TAF5 holbrookia 40 51.3

817 Q9TAG1 holbrookia 40 51.3

818 Q9TAG5 uma inornat 40 51.3

819 Q6TQV0 liolaemus o 40 51.3

820 Q9T454 uma scopari 40 51.3

821 Q9T4Q0 callisaur 40 51.3

822 Q9T3Y7 uma notata 40 51.3

823 Q6TR09 liolaemus d 40 51.3

824 Q6TQV4 liolaemus g 40 51.3

825 Q4VTV8 gymnodactyl 40 51.3

826 Q6KCT8 micoureus d 40 51.3

827 Q9MM61 eremias vel 40 51.3

828 Q6TQX8 liolaemus s 40 51.3

829 Q6RBY2 marmota mon 40 51.3

830 Q9MIB1 mabuya quin 40 51.3

831 Q85NU8 pachydactyl 40 51.3

832 Q85H40 pachydactyl 40 51.3

833 Q6TR56 liolaemus d 40 51.3

834 Q4VTV5 gymnodactyl 40 51.3

Q69fu7 arborimus l

Q69fs8 arborimus l

Q69fr6 arborimus a

Q69fr3 phenacomys

Q69fr2 phenacomys

Q69fri arborimus p

Q69fw5 arborimus l

Q69fv1 arborimus l

Q69fu2 arborimus l

Q69fu1 arborimus l

Q69fg9 arborimus p

Q69fv9 arborimus l

Q69fr8 arborimus a

Q957i9 lacerta agi

Q957i8 lacerta bil

Q8m2d4 pachydactyl

Q8m2e4 pachydactyl

Q8m2e1 pachydactyl

Q952g5 lacerta sch

Q952g3 lacerta sch

Q8m2d2 pachydactyl

Q952g4 lacerta sch

Q8m2e2 pachydactyl

Q6tcy5 liolaemus l

Q94pe2 lacerta sch

Q94pe1 lacerta sch

Q94pe0 lacerta sch

Q94pk3 lacerta sch

Q6tr08 liolaemus d

Q6tcu8 liolaemus u

Q4vtw7 gymnodactyl

Q4vtw4 gymnodactyl

Q9xmj9 brachylagus

Q9the7 sylvilagus

Q9xm1 lepus calif

Q9xm2 lepus capen

Q9xm5 sylvilagus

Q9xm0 bunolagus m

Q9the6 lepus ameri

Q9xm4 romerolagus

Q9tgr9 lepus timid

Q85i90 liolaemus e

Q7yaz3 cordylosaur

Q52n21 mabuya macu

Q52n20 mabuya macu

Q52n05 mabuya macu

Q4vtv9 gymnodactyl

Q71fd4 barbastella

Q71fc6 barbastella

Q71fc5 barbastella

Q71fb1 plecotus au

Q71fa8 plecotus au

Q71fa4 plecotus au

Q71fa3 plecotus au

Q71fa2 plecotus au

Q71fa0 plecotus cf

Q71f99 plecotus cf

Q71f97 plecotus ko

Q71f96 plecotus au

Q71f93 plecotus sp

Q71f92 plecotus sp

Q71f89 plecotus ba

Q71f88 plecotus ba

Q71f80 plecotus te

Q71f79 plecotus te

Q6w818 plecotus sp

Q71fa5 plecotus sp

Q71f91 plecotus au

Q71fd2 barbastella

Q71fc7 barbastella

Q71fa6 plecotus au

Q71f90 plecotus au

Q85te0 cavia porce

835 40 51.3 226 2 Q85TD8_CAVAP Q85td8 cavia apere 908 40 51.3 233 2 Q6TR48_9SAUR Q6tr48 liolaemus d

836 40 51.3 226 2 Q85TD7_CAVAP Q85td7 cavia apere 909 40 51.3 233 2 Q6TR25_9SAUR Q6tr25 liolaemus d

837 40 51.3 226 2 Q85TD6_9HYST Q85td6 microcavia 910 40 51.3 233 2 Q5GI03_9SAUR Q5gi03 liolaemus m

838 40 51.3 226 2 Q85TD5_9HYST Q85td5 microcavia 911 40 51.3 233 2 Q5GI02_9SAUR Q5gi02 liolaemus m

839 40 51.3 226 2 Q85TD4_GALMU Q85td4 galea muste 912 40 51.3 233 2 Q4VTY2_9SAUR Q4vty2 gymnodactyl

840 40 51.3 226 2 Q6TR45_9SAUR Q6tr45 liolaemus d 913 40 51.3 233 2 Q4VTW3_9SAUR Q4vtw3 gymnodactyl

841 40 51.3 226 2 Q6TR36_9SAUR Q6tr36 liolaemus d 914 40 51.3 233 2 Q4VTW8_9SAUR Q4vtw8 gymnodactyl

842 40 51.3 226 2 Q6TQY9_9SAUR Q6tqy9 liolaemus l 915 40 51.3 233 2 Q4VTW3_9SAUR Q4vtw3 gymnodactyl

843 40 51.3 226 2 Q52N07_9SAUR Q52n07 mabuya macu 916 40 51.3 234 1 CYB_LEPAL 047558 lepus allen

844 40 51.3 226 2 Q4VTW9_9SAUR Q4vtw9 gymnodactyl 917 40 51.3 234 1 CYB_LEPAR 047555 lepus arcti

845 40 51.3 227 2 Q9B7W2_9CETA Q9b7w2 platanista 918 40 51.3 234 1 CYB_LEPOT 047556 lepus othus

846 40 51.3 227 2 Q6TR66_9SAUR Q6tr66 liolaemus d 919 40 51.3 234 2 Q9XKP5_LEPTO Q9xkp5 lepus town

847 40 51.3 227 2 Q6TR66_9SAUR Q6tr66 liolaemus d 920 40 51.3 234 2 Q9XKP4_LEPSX Q9xkp4 lepus saxat

848 40 51.3 227 2 Q6TQX1_9SAUR Q6tqx1 liolaemus g 921 40 51.3 234 2 Q7YAU9_TAMRU Q7yau9 tamias rufi

849 40 51.3 227 2 Q52N12_9SAUR Q52n12 mabuya macu 922 40 51.3 234 2 Q7Y6Y4_TAMAM Q7y6y4 tamias amoe

850 40 51.3 227 2 Q52N04_9SAUR Q52n04 mabuya macu 923 40 51.3 234 2 Q7Y798_TAMRU Q7y798 tamias amoe

851 40 51.3 227 2 Q4VTV7_9SAUR Q4vtv7 gymnodactyl 924 40 51.3 234 2 Q7Y799_9SCIU Q7y799 tamias amoe

852 40 51.3 228 2 Q9MHR7_9SAUR Q9mhr7 pachydactyl 925 40 51.3 234 2 Q7Y797_TAMAM Q7y797 tamias amoe

853 40 51.3 228 2 Q9MHR4_9SAUR Q9mhr4 tarentola b 926 40 51.3 234 2 Q7Y7J7_TAMAM Q7y7j7 tamias amoe

854 40 51.3 228 2 Q9MHR6_9SAUR Q9mhr6 tarentola a 927 40 51.3 234 2 Q7Y721_TAMAM Q7y721 tamias amoe

855 40 51.3 228 2 Q52N32_9SAUR Q52n32 mabuya macu 928 40 51.3 234 2 Q85IB5_9SAUR Q85ib5 liolaemus p

856 40 51.3 228 2 Q52N24_9SAUR Q52n24 mabuya macu 929 40 51.3 234 2 Q85IB5_9SAUR Q85ib5 liolaemus p

857 40 51.3 228 2 Q52N18_9SAUR Q52n18 mabuya macu 930 40 51.3 234 2 Q85I73_9SAUR Q85i73 liolaemus p

858 40 51.3 228 2 Q52N17_9SAUR Q52n17 mabuya macu 931 40 51.3 234 2 Q85I72_9SAUR Q85i72 liolaemus e

859 40 51.3 228 2 Q52N13_9SAUR Q52n13 mabuya macu 932 40 51.3 234 2 Q6TR64_9SAUR Q6tr64 liolaemus d

860 40 51.3 228 2 Q52N08_9SAUR Q52n08 mabuya macu 933 40 51.3 234 2 Q6TR60_9SAUR Q6tr60 liolaemus d

861 40 51.3 228 2 Q52N44_9SAUR Q52n44 mabuya macu 934 40 51.3 234 2 Q6TR26_9SAUR Q6tr26 liolaemus d

862 40 51.3 228 2 Q52N34_9SAUR Q52n34 mabuya macu 935 40 51.3 234 2 Q4VTX7_9SAUR Q4vtx7 gymnodactyl

863 40 51.3 228 2 Q52N25_9SAUR Q52n25 mabuya macu 936 40 51.3 235 2 Q85DK1_APOSY Q85dk1 apodemus sy

864 40 51.3 228 2 Q52N33_9SAUR Q52n33 mabuya macu 937 40 51.3 235 2 Q85IA5_9SAUR Q85ias liolaemus e

865 40 51.3 228 2 Q4VTW5_9SAUR Q4vtw5 gymnodactyl 938 40 51.3 235 2 Q85I96_9SAUR Q85i96 liolaemus s

866 40 51.3 229 2 Q70EG3_9RODE Q70eg3 saccostomus 939 40 51.3 235 2 Q85I95_9SAUR Q85i95 liolaemus s

867 40 51.3 229 2 Q9G2I7_9RODE Q9g2i7 peromyscus 940 40 51.3 235 2 Q85I93_9SAUR Q85i93 liolaemus s

868 40 51.3 229 2 Q6TR02_9SAUR Q6tr02 liolaemus d 941 40 51.3 235 2 Q85I89_9SAUR Q85i89 liolaemus e

869 40 51.3 229 2 Q6TQY6_9SAUR Q6tqy6 liolaemus l 942 40 51.3 235 2 Q85I87_9SAUR Q85i87 liolaemus s

870 40 51.3 230 2 Q6TR31_9SAUR Q6tr31 liolaemus d 943 40 51.3 235 2 Q85I80_9SAUR Q85i80 liolaemus s

871 40 51.3 230 2 Q85IA8_9SAUR Q85ia8 liolaemus k 944 40 51.3 235 2 Q85I77_9SAUR Q85i77 liolaemus e

872 40 51.3 230 2 Q85IA2_9SAUR Q85ia2 liolaemus e 945 40 51.3 235 2 Q85I83_9SAUR Q85i83 liolaemus s

873 40 51.3 230 2 Q85I86_9SAUR Q85i86 liolaemus s 946 40 51.3 235 2 Q6TR62_9SAUR Q6tr62 liolaemus d

874 40 51.3 230 2 Q7YAX0_9SAUR Q7yax0 typhlaconti 947 40 51.3 235 2 Q6TR21_9SAUR Q6tr21 liolaemus d

875 40 51.3 230 2 Q6TR59_9SAUR Q6tr59 liolaemus d 948 40 51.3 235 2 Q85BP5_9SAUR Q85bp5 liolaemus e

876 40 51.3 230 2 Q6TR06_9SAUR Q6tr06 liolaemus d 949 40 51.3 235 2 Q85CJ4_9SAUR Q85cj4 liolaemus k

877 40 51.3 230 2 Q6TQX9_9SAUR Q6tqx9 liolaemus s 950 40 51.3 235 2 Q4VTY1_9SAUR Q4vty1 gymnodactyl

878 40 51.3 230 2 Q6TQW0_9SAUR Q6tqw0 liolaemus g 951 40 51.3 235 2 Q4VTX5_9SAUR Q4vtx5 gymnodactyl

879 40 51.3 230 2 Q6TQZ1_9SAUR Q6tqz1 liolaemus l 952 40 51.3 236 2 Q4VTV0_9SAUR Q4vtv0 gymnodactyl

880 40 51.3 230 2 Q6TR23_9SAUR Q6tr23 liolaemus d 953 40 51.3 236 2 Q8WCG3_TAMSE Q8wcg3 tamias sene

881 40 51.3 230 2 Q4VTX1_9SAUR Q4vtx1 gymnodactyl 954 40 51.3 236 2 Q8WCG4_TAMRU Q8wcg4 tamias rufi

882 40 51.3 230 2 Q4VTU9_9SAUR Q4vtu9 gymnodactyl 955 40 51.3 236 2 Q8WCG0_TAMAM Q8wcg0 tamias amoe

883 40 51.3 231 2 Q6TR44_9SAUR Q6tr44 liolaemus d 956 40 51.3 236 2 Q8WCG1_TAMMI Q8wcg1 tamias mini

884 40 51.3 231 2 Q6TR42_9SAUR Q6tr42 liolaemus d 957 40 51.3 236 2 Q8M7F3_9RODE Q8m7f3 zygodontomy

885 40 51.3 231 2 Q6TQU9_9SAUR Q6tqu9 liolaemus c 958 40 51.3 236 2 Q8WCG2_TAMTO Q8wcg2 tamias town

886 40 51.3 231 2 Q6TQU7_9SAUR Q6tqu7 liolaemus k 959 40 51.3 236 2 Q8W988_TAMRU Q8w988 tamias rufi

887 40 51.3 231 2 Q6TQU6_9SAUR Q6tqu6 liolaemus a 960 40 51.3 236 2 Q8W7M5_TAMRU Q8w7m5 tamias rufi

888 40 51.3 231 2 Q6TQU4_9SAUR Q6tqu4 liolaemus b 961 40 51.3 236 2 Q8W8A1_TAMRU Q8w8a1 tamias rufi

889 40 51.3 231 2 Q85CU5_9SAUR Q85cu5 liolaemus b 962 40 51.3 236 2 Q8W7C5_TAMRU Q8w7c5 tamias rufi

890 40 51.3 231 2 Q4VTV3_9SAUR Q4vtv3 gymnodactyl 963 40 51.3 236 2 Q6TQZ9_9SAUR Q6tqz9 liolaemus d

891 40 51.3 232 2 Q85IA0_9SAUR Q85ia0 liolaemus k 964 40 51.3 237 2 Q4JFM2_CYNME Q4jfm2 cynomys mex

892 40 51.3 232 2 Q85IA9_9SAUR Q85ia9 liolaemus s 965 40 51.3 237 2 Q4JFM3_CYNME Q4jfm3 cynomys mex

893 40 51.3 232 2 Q85I84_9SAUR Q85i84 liolaemus s 966 40 51.3 237 2 Q4LBD1_9MURI Q4lbd1 mus matthey

894 40 51.3 232 2 Q85I92_9SAUR Q85i92 liolaemus s 967 40 51.3 237 2 Q8M7F5_9RODE Q8m7f5 juliomys ri

895 40 51.3 232 2 Q6TR32_9SAUR Q6tr32 liolaemus d 968 40 51.3 237 2 Q8WEC2_9SAUR Q8wec2 mabuya marg

896 40 51.3 232 2 Q4VTX8_9SAUR Q4vtx8 gymnodactyl 969 40 51.3 237 2 Q8WED8_9SAUR Q8wed8 cryptobleph

897 40 51.3 232 2 Q4VTW1_9SAUR Q4vtw1 gymnodactyl 970 40 51.3 237 2 Q8WEC1_9SAUR Q8wec1 mabuya macu

898 40 51.3 233 2 Q9T6J3_LEPCB Q9t6j3 lepus comus 971 40 51.3 237 2 Q8WEC0_9SAUR Q8wec0 mabuya macu

899 40 51.3 233 2 Q85Q52_9HYST Q85q52 thrichomys 972 40 51.3 237 2 Q9MM67_9SAUR Q9mm67 mesalina br

900 40 51.3 233 2 Q7YB00_GAMWI Q7yb00 gambelia wi 973 40 51.3 237 2 Q8SK11_9SAUR Q8sk11 amphiglossu

901 40 51.3 233 2 Q7YAY6_9SAUR Q7yay6 emoia jakat 974 40 51.3 237 2 Q85IB2_9SAUR Q85ib2 liolaemus p

902 40 51.3 233 2 Q7YAY0_9SAUR Q7yay0 lygisaurus 975 40 51.3 237 2 Q6XNT0_9SAUR Q6xnt0 phelsuma gu

903 40 51.3 233 2 Q7YAX9_9SAUR Q7yax9 tiliqua gig 976 40 51.3 237 2 Q6TQZ6_9SAUR Q6tqz6 liolaemus d

904 40 51.3 233 2 Q7YAX6_EMOCY Q7yax6 emoia cyanu 977 40 51.3 237 2 Q6TQY3_9SAUR Q6tqy3 liolaemus l

905 40 51.3 233 2 Q7YAX5_9SAUR Q7yax5 mesalina gu 978 40 51.3 237 2 Q6TQV2_9SAUR Q6tqv2 liolaemus p

906 40 51.3 233 2 Q7YAW7_9SAUR Q7yaw7 typhlaconti 979 40 51.3 237 2 Q85C60_9SAUR Q85c60 liolaemus p

907 40 51.3 233 2 Q7YAW5_9SAUR Q7yaw5 scelotes bi 980 40 51.3 237 2 Q58VU7_9SAUR Q58vu7 cryptobleph

908 40 51.3 233 2 Q6TR48_9SAUR Q6tr48 liolaemus d

909 40 51.3 233 2 Q6TR25_9SAUR Q6tr25 liolaemus d

910 40 51.3 233 2 Q5GI03_9SAUR Q5gi03 liolaemus m

911 40 51.3 233 2 Q5GI02_9SAUR Q5gi02 liolaemus m

912 40 51.3 233 2 Q4VTY2_9SAUR Q4vty2 gymnodactyl

913 40 51.3 233 2 Q4VTW3_9SAUR Q4vtw3 gymnodactyl

914 40 51.3 233 2 Q4VTW8_9SAUR Q4vtw8 gymnodactyl

915 40 51.3 233 2 Q4VTW3_9SAUR Q4vtw3 gymnodactyl

916 40 51.3 234 1 CYB_LEPAL 047558 lepus allen

917 40 51.3 234 1 CYB_LEPAR 047555 lepus arcti

918 40 51.3 234 1 CYB_LEPOT 047556 lepus othus

919 40 51.3 234 2 Q9XKP5_LEPTO Q9xkp5 lepus town

920 40 51.3 234 2 Q9XKP4_LEPSX Q9xkp4 lepus saxat

921 40 51.3 234 2 Q7YAU9_TAMRU Q7yau9 tamias rufi

922 40 51.3 234 2 Q7Y6Y4_TAMAM Q7y6y4 tamias amoe

923 40 51.3 234 2 Q7Y798_TAMRU Q7y798 tamias amoe

924 40 51.3 234 2 Q7Y799_9SCIU Q7y799 tamias amoe

925 40 51.3 234 2 Q7Y797_TAMAM Q7y797 tamias amoe

926 40 51.3 234 2 Q7Y7J7_TAMAM Q7y7j7 tamias amoe

927 40 51.3 234 2 Q7Y721_TAMAM Q7y721 tamias amoe

928 40 51.3 234 2 Q85IB5_9SAUR Q85ib5 liolaemus p

929 40 51.3 234 2 Q85IB5_9SAUR Q85ib5 liolaemus p

930 40 51.3 234 2 Q85I73_9SAUR Q85i73 liolaemus p

931 40 51.3 234 2 Q85I72_9SAUR Q85i72 liolaemus e

932 40 51.3 234 2 Q6TR64_9SAUR Q6tr64 liolaemus d

933 40 51.3 234 2 Q6TR60_9SAUR Q6tr60 liolaemus d

934 40 51.3 234 2 Q6TR26_9SAUR Q6tr26 liolaemus d

935 40 51.3 234 2 Q4VTX7_9SAUR Q4vtx7 gymnodactyl

936 40 51.3 235 2 Q85DK1_APOSY Q85dk1 apodemus sy

937 40 51.3 235 2 Q85IA5_9SAUR Q85ias liolaemus e

938 40 51.3 235 2 Q85I96_9SAUR Q85i96 liolaemus s

939 40 51.3 235 2 Q85I95_9SAUR Q85i95 liolaemus s

940 40 51.3 235 2 Q85I93_9SAUR Q85i93 liolaemus s

941 40 51.3 235 2 Q85I89_9SAUR Q85i89 liolaemus e

942 40 51.3 235 2 Q85I87_9SAUR Q85i87 liolaemus s

943 40 51.3 235 2 Q85I80_9SAUR Q85i80 liolaemus s

944 40 51.3 235 2 Q85I77_9SAUR Q85i77 liolaemus e

945 40 51.3 235 2 Q85I83_9SAUR Q85i83 liolaemus s

946 40 51.3 235 2 Q6TR62_9SAUR Q6tr62 liolaemus d

947 40 51.3 235 2 Q6TR21_9SAUR Q6tr21 liolaemus d

948 40 51.3 235 2 Q85BP5_9SAUR Q85bp5 liolaemus e

949 40 51.3 235 2 Q85CJ4_9SAUR Q85cj4 liolaemus k

950 40 51.3 235 2 Q4VTY1_9SAUR Q4vty1 gymnodactyl

951 40 51.3 235 2 Q4VTX5_9SAUR Q4vtx5 gymnodactyl

952 40 51.3 236 2 Q4VTV0_9SAUR Q4vtv0 gymnodactyl

953 40 51.3 236 2 Q8WCG3_TAMSE Q8wcg3 tamias sene

954 40 51.3 236 2 Q8WCG4_TAMRU Q8wcg4 tamias rufi

955 40 51.3 236 2 Q8WCG0_TAMAM Q8wcg0 tamias amoe

956 40 51.3 236 2 Q8WCG1_TAMMI Q8wcg1 tamias mini

957 40 51.3 236 2 Q8M7F3_9RODE Q8m7f3 zygodontomy

958 40 51.3 236 2 Q8WCG2_TAMTO Q8wcg2 tamias town

959 40 51.3 236 2 Q8W988_TAMRU Q8w988 tamias rufi

960 40 51.3 236 2 Q8W7M5_TAMRU Q8w7m5 tamias rufi

961 40 51.3 236 2 Q8W8A1_TAMRU Q8w8a1 tamias rufi

962 40 51.3 236 2 Q8W7C5_TAMRU Q8w7c5 tamias rufi

963 40 51.3 236 2 Q6TQZ9_9SAUR Q6tqz9 liolaemus d

964 40 51.3 237 2 Q4JFM2_CYNME Q4jfm2 cynomys mex

965 40 51.3 237 2 Q4JFM3_CYNME Q4jfm3 cynomys mex

966 40 51.3 237 2 Q4LBD1_9MURI Q4lbd1 mus matthey

967 40 51.3 237 2 Q8M7F5_9RODE Q8m7f5 juliomys ri

968 40 51.3 237 2 Q8WEC2_9SAUR Q8wec2 mabuya marg

969 40 51.3 237 2 Q8WED8_9SAUR Q8wed8 cryptobleph

970 40 51.3 237 2 Q8WEC1_9SAUR Q8wec1 mabuya macu

971 40 51.3 237 2 Q8WEC0_9SAUR Q8wec0 mabuya macu

972 40 51.3 237 2 Q9MM67_9SAUR Q9mm67 mesalina br

973 40 51.3 237 2 Q8SK11_9SAUR Q8sk11 amphiglossu

974 40 51.3 237 2 Q85IB2_9SAUR Q85ib2 liolaemus p

975 40 51.3 237 2 Q6XNT0_9SAUR Q6xnt0 phelsuma gu

976 40 51.3 237 2 Q6TQZ6_9SAUR Q6tqz6 liolaemus d

977 40 51.3 237 2 Q6TQY3_9SAUR Q6tqy3 liolaemus l

978 40 51.3 237 2 Q6TQV2_9SAUR Q6tqv2 liolaemus p

979 40 51.3 237 2 Q85C60_9SAUR Q85c60 liolaemus p

980 40 51.3 237 2 Q58VU7_9SAUR Q58vu7 cryptobleph

981 40 51.3 238 2 Q9T6I6_LEPOI Q9t6i6 lepus oiost
982 40 51.3 238 2 Q4JFM0_SPESP Q4jfm0 spermophilu
983 40 51.3 238 2 Q4JFM1_SPESP Q4jfm1 spermophilu
984 40 51.3 238 2 O03375_9SCIU O03375 paraxerus c
985 40 51.3 238 2 Q6XNZ7_9SAUR Q6xnz7 pachydactyl
986 40 51.3 238 2 Q6XNZ6_9SAUR Q6xnz6 rhoptropus
987 40 51.3 238 2 Q6XNZ3_9SAUR Q6xnz3 phelsuma mu
988 40 51.3 238 2 Q6XNZ2_9SAUR Q6xnz2 phelsuma as
989 40 51.3 238 2 Q6XNY9_9SAUR Q6xny9 phelsuma se
990 40 51.3 238 2 Q6XNY7_9SAUR Q6xny7 phelsuma li
991 40 51.3 238 2 Q6XNY6_9SAUR Q6xny6 phelsuma st
992 40 51.3 238 2 Q6XNY4_9SAUR Q6xny4 phelsuma qu
993 40 51.3 238 2 Q6XNY3_9SAUR Q6xny3 phelsuma ab
994 40 51.3 238 2 Q6XNY2_9SAUR Q6xny2 phelsuma ab
995 40 51.3 238 2 Q6XNY1_9SAUR Q6xny1 phelsuma ab
996 40 51.3 238 2 Q6XNY0_9SAUR Q6xny0 phelsuma ab
997 40 51.3 238 2 Q6XNS9_9SAUR Q6xns9 phelsuma gu
998 40 51.3 238 2 Q6UKV6_9SAUR Q6ukv6 phelaemus b
999 40 51.3 238 2 Q6TR69_9SAUR Q6tr69 liolaemus d
1000 40 51.3 238 2 Q6TR63_9SAUR Q6tr63 liolaemus d

ALIGNMENTS

RESULT 1
CX6A2_BOVIN
ID CX6A2_BOVIN STANDARD; PRT; 97 AA.
AC P07471;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Cytochrome c oxidase polypeptide Via-heart, mitochondrial precursor
DE (EC 1.9.3.1) (COXVIAH) (Polypeptide Vlb).
GN Name=COX6A2; Synonyms=COX6A;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=91274363; PubMed=1647214; DOI=10.1016/0167-4781(91)90022-E;
RA Smith E.O., Bement D.M., Grossman L.I., Lomax M.I.;
RT "The cDNA for the heart/muscle isoform of bovine cytochrome c oxidase subunit Via encodes a presequence.";
RL Biochim. Biophys. Acta 1089:266-268(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93326637; PubMed=7687470; DOI=10.1016/0167-4781(93)90092-R;
RA Smith E.O., Lomax M.I.;
RT "Structural organization of the bovine gene for the heart/muscle isoform of cytochrome c oxidase subunit Via.";
RL Biochim. Biophys. Acta 1174:63-71(1993).
RN [3]
RP PROTEIN SEQUENCE OF 13-96.
RC TISSUE=Heart;
RX MEDLINE=86000134; PubMed=2994692;
RA Meinecke L., Buse G.;
RT "Studies on cytochrome c oxidase, XII. Isolation and primary structure of polypeptide Vlb from bovine heart.";
RL Biol. Chem. Hoppe-Seyler 366:687-694(1985).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=96216288; PubMed=8638158;
RA Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H., Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
RT "The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A.";
RL Science 272:1136-1144(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RC TISSUE=Heart;
RX MEDLINE=99190827; PubMed=10089392;
RA Tomizaki T., Yamashita E., Yamaguchi H., Aoyama H., Tsukihara T., Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
RT "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A resolution.";
RL Acta Crystallogr. D 55:31-45(1999).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Heart;
RX MEDLINE=20235645; PubMed=10771420; DOI=10.1107/S0907444900002213;
RA Fei M.J., Yamashita E., Inoue N., Yao M., Yamaguchi H., Tsukihara T., Shinzawa-Itoh K., Nakashima R., Yoshikawa S.;
RT "X-ray structure of azide-bound fully oxidized cytochrome c oxidase from bovine heart at 2.9 A resolution.";
RL Acta Crystallogr. D 56:529-535(2000).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide chains of cytochrome c oxidase, the terminal oxidase in mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- TISSUE SPECIFICITY: Heart/muscle specific isoform.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase Via family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
CC EMBL; S64127; AAB27605.1; -; Genomic_DNA.
DR EMBL; X56857; CAA40183.1; -; mRNA.
DR PIR; S35702; OGBO6.
DR PDB; 1OCC; X-ray; G/T=13-96.
DR PDB; 1OCO; X-ray; G/T=13-96.
DR PDB; 1OCR; X-ray; G/T=13-96.
DR PDB; 1OCZ; X-ray; G/T=13-96.
DR PDB; 1V54; X-ray; G/T=13-97.
DR PDB; 1V55; X-ray; G/T=13-97.
DR PDB; 2OCC; X-ray; G/T=13-96.
DR InterPro; IPR001349; COX6A.
DR PANTHER; PTHR11504; COX6A; 1.
DR Pfam; PF02046; COX6A; 1.
DR PIRSF; PIRSF000277; COX6A1; 1.
DR ProDom; PD006036; COX6A; 1.
DR PROSITE; PS01329; COX6A; 1.
KW 3D-structure; Direct protein sequencing; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Transit peptide; Transmembrane.
FT TRANSIT 1 12 Mitochondrion.
FT CHAIN 13 97 Cytochrome c oxidase polypeptide Via-heart.
FT FT Mitochondrial matrix.
FT TOPO_DOM 13 24 Mitochondrial matrix.
FT TRANSMEM 25 49 Mitochondrial intermembrane.
FT TOPO_DOM 50 97 Mitochondrial intermembrane.
SQ SEQUENCE 97 AA; 10800 MW; 59B408A1B9B6AF4F CRC64;

Query Match 60.3%; Score 47; DB 1; Length 97;
Best Local Similarity 70.0%; Pred. NO. 7.6;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGG 10
||| ||| |
Db 43 CTLSNWLHSG 52

RESULT 2
Q9GAB0_9BIVA
ID Q9GAB0_9BIVA PRELIMINARY; PRT; 128 AA.
AC Q9GAB0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Cytochrome b (Fragment).
GN Name=cytb;
OS Pitar morrhuanus (false quahog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Pitar.
OX NCBI_TaxID=115818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-cl complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205082; AAG32644.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 13986 MW; EA78F42E59583BEE CRC64;

Query Match 60.3%; Score 47; DB 2; Length 128;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
Db :|||||
18 SLVEWLWGGFC 28

RESULT 3
Q9GAB2_MERCA
ID Q9GAB2_MERCA PRELIMINARY; PRT; 128 AA.
AC Q9GAB2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Mercenaria campechiensis (Southern quahog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Mercenaria.
OX NCBI_TaxID=115816;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-cl complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205081; AAG32643.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; Cytochrom_B_N; 1.
DR PROSITE; PS51002; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 13986 MW; EA78F42E59583BEE CRC64;

CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205080; AAG32642.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14058 MW; 2B3C1546BBC9B8F3 CRC64;

Query Match 59.0%; Score 46; DB 2; Length 128;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LREWLHGGFC 12
Db :|||||
19 LVEWLWGGFC 28

RESULT 4
Q9GAB1_9BIVA
ID Q9GAB1_9BIVA PRELIMINARY; PRT; 128 AA.
AC Q9GAB1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Calyptogena sp. 'Kodiak Seep'.
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Glossoida; Vesicomysidae; Calyptogena.
OX NCBI_TaxID=118660;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-cl complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205081; AAG32643.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.

KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14436 MW; CDA42D1A562B0400 CRC64;

Query Match 59.0%; Score 46; DB 2; Length 128;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LREWLHGGFC 12
Db 19 LVEWLWGGFC 28

RESULT 5
Q6H5Y4 ORYSA PRELIMINARY; PRT; 279 AA.
AC Q6H5Y4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative receptor-like kinase.
GN Name=P0620H05.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005394; BAD25865.1; -; Genomic_DNA.
DR Gramine; Q6H5Y4; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Receptor.
SQ SEQUENCE 279 AA; 31269 MW; F201A2700EFP5P2BF CRC64;

Query Match 59.0%; Score 46; DB 2; Length 279;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
Db 86 CCLHEWLHG 94

RESULT 6
Q75D31 ASHGO PRELIMINARY; PRT; 462 AA.
AC Q75D31;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ABR192Wp.
GN Name=ABR192W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavler A.,
RA Gafney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016815; AAS50964.1; -; Genomic_DNA.
DR AGD; ABR192W; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; tRNA-synt_lb.
DR InterPro; IPR002307; Tyr tRNA-synt_lb.
DR Pfam; PF00579; tRNA-synt_lb; 1.
DR PRINTS; PR01040; TRNASYNTHTYR.
DR TIGRFAMs; TIGR00234; tyrS; 1.
KW Complete proteome.
SQ SEQUENCE 462 AA; 50532 MW; C6D4D33BCAEA9FC1 CRC64;

Query Match 59.0%; Score 46; DB 2; Length 462;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLREWLHGG 10
Db 7 TMRPWLHGG 15

RESULT 7
Q4XYI2 PLACH PRELIMINARY; PRT; 767 AA.
ID Q4XYI2 PLACH PRELIMINARY;
AC Q4XYI2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC000713.02.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAJ01002255; CAH78028.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON_TER 1 767
FT NON_TER 767 767
SQ SEQUENCE 767 AA; 88785 MW; C82CEDC360481D9B CRC64;

Query Match 59.0%; Score 46; DB 2; Length 767;
Best Local Similarity 63.6%; Pred. No. 90;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 69 CTLQKWLFGAF 79

```

RESULT 8
Q4YVE2_PLABE
ID Q4YVE2_PLABE PRELIMINARY; PRT; 1429 AA.
AC Q4YVE2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PB000699.02.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RA "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAI0100277; CAH98014.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1429 AA; 167259 MW; DF23746492FF1DE9 CRC64;

Query Match 59.0%; Score 46; DB 2; Length 1429;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 180 CTLQKWLFGAF 190

RESULT 9
Q8IKF6_PLAF7
ID Q8IKF6_PLAF7 PRELIMINARY; PRT; 2558 AA.
AC Q8IKF6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PF14_0649;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014827; AAN37262.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.

```

```

DR InterPro; IPR011498; Kelch_2.
DR Pfam; PF07646; Kelch_2; 2.
DR SMART; SM00387; HATPase_C; 1.
KW Hypothetical protein.
SQ SEQUENCE 2558 AA; 295828 MW; 6B00A5209A3489AB CRC64;

Query Match 59.0%; Score 46; DB 2; Length 2558;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 1282 CTLQKWLFGAF 1292

RESULT 10
Q7PDT6_PLAYO
ID Q7PDT6_PLAYO PRELIMINARY; PRT; 2584 AA.
AC Q7PDT6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte membrane protein PFEMP3.
GN Name=PY01412;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000372; EAA20743.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR Pfam; PF02518; HATPase_C; 1.
SQ SEQUENCE 2584 AA; 297199 MW; D5E87355F355007E CRC64;

Query Match 59.0%; Score 46; DB 2; Length 2584;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 1283 CTLQKWLFGAF 1293

RESULT 11
Q8TJEO_METAC
ID Q8TJEO_METAC PRELIMINARY; PRT; 126 AA.
AC Q8TJEO;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=MA3845;
OS Methanosarcina acetivorans.

```

```
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C2A / ATCC 35395 / DSM 2834; DOI=10.1101/gr.223902;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeAtellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AB011095; AAM07196.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 126 AA; 14491 MW; 30FC5B8FDEA11926 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 126;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHGG 10
||:|:|:|
DB 70 CTIREYLHDG 79

RESULT 12
Q9ZNB4_PORGI
ID Q9ZNB4_PORGI PRELIMINARY; PRT; 211 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Insertion sequence IS1126-like gene.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RX MEDLINE=99143166; PubMed=9988746; DOI=10.1074/jbc.274.8.5012;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products.";
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL; AB019363; BAA34342.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 211 AA; 24444 MW; 77040F65D76D1175 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
::|:|:|:|
DB 167 SRRWFHGGRC 177

RESULT 13
Q9ZNB4_PORGI
ID Q9ZNB4_PORGI PRELIMINARY; PRT; 211 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Insertion sequence IS1126-like gene.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RX MEDLINE=99143166; PubMed=9988746; DOI=10.1074/jbc.274.8.5012;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products.";
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL; AB019363; BAA34342.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 211 AA; 24444 MW; 77040F65D76D1175 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
::|:|:|:|
DB 167 SRRWFHGGRC 177
```

```
Q7MUZ9_PORGI
ID Q7MUZ9_PORGI PRELIMINARY; PRT; 211 AA.
AC Q7MUZ9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ISPg1, transposase, internal deletion.
GN OrderedLocusNames=PGI320;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017176; AAQ66391.1; -; Genomic_DNA.
DR TIGR; PG1320; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Complete proteome.
SQ SEQUENCE 211 AA; 24300 MW; 86F8E3CAA3289A71 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
::|:|:|:|
DB 167 SRRWFHGGRC 177

RESULT 14
Q9ZA61_PORGI
ID Q9ZA61_PORGI PRELIMINARY; PRT; 361 AA.
AC Q9ZA61;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=99150246; PubMed=10024556;
RA Hanley S.A., Aduse-Opoku J., Curtis M.A.;
RT "A 55-kilodalton immunodominant antigen of Porphyromonas gingivalis
RT W50 has arisen via horizontal gene transfer.";
RL Infect. Immun. 67:1157-1171(1999).
DR EMBL; AJ130872; CAA1025.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 361 AA; 41821 MW; F43664DC40EE012C CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
```


Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
:| | | | |
Db 317 SRRWFHGGRC 327

RESULT 15

Q9ZAD0_PORGI
ID Q9ZAD0_PORGI PRELIMINARY; PRT; 361 AA.
AC Q9ZAD0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF7.
GN Name=orf7;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OS Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33277;
RX MEDLINE=99198959; PubMed=10100860; DOI=10.1016/S0014-5793(99)00237-9;
RA Yoshida A., Nakano Y., Yamashita Y., Oho T., Shibata Y., Ohishi M.,
RA Koga T.;
RT "A novel dnaK operon from Porphyromonas gingivalis.";
RL FEBS Lett. 446:287-291(1999).
DR EMBL; AB015879; BAA35090.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 361 AA; 41588 MW; F47766D275D597F1 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
:| | | | |
Db 317 SRRWFHGGRC 327

RESULT 16

Q7M7B5_PORGI
ID Q7M7B5_PORGI PRELIMINARY; PRT; 361 AA.
AC Q7M7B5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE ISPg1, transposase.
GN OrderedLocusNames=PG0184, PG0825, PG1031, PG1177, PG1197, PG1448,
GN PG1624, PG1906;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).

DR EMBL; AE017172; AAQ65419.1; -; Genomic_DNA.
DR EMBL; AE017174; AAQ65980.1; -; Genomic_DNA.
DR EMBL; AE017175; AAQ66150.1; -; Genomic_DNA.
DR EMBL; AE017176; AAQ66277.1; -; Genomic_DNA.
DR EMBL; AE017176; AAQ66289.1; -; Genomic_DNA.
DR EMBL; AE017177; AAQ66497.1; -; Genomic_DNA.
DR EMBL; AE017177; AAQ66552.1; -; Genomic_DNA.
DR EMBL; AE017178; AAQ66889.1; -; Genomic_DNA.
DR TIGR; PG0184; -.
DR TIGR; PG0825; -.
DR TIGR; PG1031; -.
DR TIGR; PG1177; -.
DR TIGR; PG1197; -.
DR TIGR; PG1448; -.
DR TIGR; PG1624; -.
DR TIGR; PG1906; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Complete proteome.
SQ SEQUENCE 361 AA; 41754 MW; 8D6D7CB8D3104A2C CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
:| | | | |
Db 317 SRRWFHGGRC 327

RESULT 17

Q7M7E9_PORGI
ID Q7M7E9_PORGI PRELIMINARY; PRT; 361 AA.
AC Q7M7E9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE ISPg1, transposase.
GN OrderedLocusNames=PG0460, PG0549, PG1244;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017173; AAQ65657.1; -; Genomic_DNA.
DR EMBL; AE017174; AAQ65741.1; -; Genomic_DNA.
DR EMBL; AE017176; AAQ66328.1; -; Genomic_DNA.
DR TIGR; PG0460; -.
DR TIGR; PG0549; -.
DR TIGR; PG1244; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Complete proteome.
SQ SEQUENCE 361 AA; 41588 MW; 394DDC1E40F7B0E1 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
Db 317 SIRRWFHGGRC 327

RESULT 18
Q67UT0_ORYSA
ID Q67UT0_ORYSA PRELIMINARY; PRT; 396 AA.
AC Q67UT0;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Putative serine/threonine protein kinase.
GN Name=P0046G12.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
clone:P0046G12.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP005419; BAD38089.1; -; Genomic_DNA.
DR Gramene; Q67UT0; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 396 AA; 43831 MW; B2C9FE83CF34445E CRC64;

Query Match 57.7%; Score 45; DB 2; Length 396;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLREWLHGG 10
Db 197 TLAQWLHGG 205

RESULT 19
Q9HRG8_HALSA
ID Q9HRG8_HALSA PRELIMINARY; PRT; 503 AA.
AC Q9HRG8;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Vng0705c.
GN OrderedLocusNames=VNG0705C;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.

OX NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AB005016; AAG19190.1; -; Genomic_DNA.
DR PIR; B84228; B84228.
DR InterPro; IPR007357; DPRP.
DR Pfam; PF04244; DPRP; 1.
KW Complete proteome.
SQ SEQUENCE 503 AA; 56619 MW; 25D5ACCFB886BC1A CRC64;

Query Match 57.7%; Score 45; DB 2; Length 503;
Best Local Similarity 77.8%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
Db 379 LTEWFHGGF 387

RESULT 20
CBPY_YEAST
ID CBPY_YEAST STANDARD; PRT; 532 AA.
AC P00729;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
GN Name=PRC1; OrderedLocusNames=YMR297W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=87131100; PubMed=3028649; DOI=10.1016/0092-8674(87)90085-7;
RA Valls L.A., Hunter C.P., Rothman J.H., Stevens T.H.;
RT "Protein sorting in yeast: the localization determinant of yeast
vacuolar carboxypeptidase Y resides in the propeptide.";
RL Cell 48:887-897(1987).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93(1997).
RN [3]
RP PROTEIN SEQUENCE OF 112-532.
RA Svendsen I., Martin B.M., Viswanatha T., Johansen J.T.;
RT "Amino acid sequence of carboxypeptidase Y. II. Peptides from
enzymatic cleavages.";
RL Carlsberg Res. Commun. 47:15-27(1982).
RN [4]
RP SEQUENCE REVISION, AND ACTIVE SITE SER-257.
RA Breddam K., Svendsen I.;
RT "Identification of methionyl and cysteinyl residues in the substrate
binding site of carboxypeptidase Y.";

RL Carlsberg Res. Commun. 49:639-645(1984).

RN [5]

RP ACTIVE SITE HIS-508.

RX MEDLINE=90315013; PubMed=2639680;

RA Bech L.M., Breddam K.;

RA "Inactivation of carboxypeptidase Y by mutational removal of the

RT putative essential histidyl residue.";

RT Carlsberg Res. Commun. 54:165-171(1989).

RL [6]

RN MUTAGENESIS.

RP MEDLINE=94114535; PubMed=7904479;

RX Mortensen U.H., Remington S.J., Breddam K.;

RA "Site-directed mutagenesis on (serine) carboxypeptidase Y. A hydrogen

RT bond network stabilizes the transition state by interaction with the

RT C-terminal carboxylate group of the substrate.";

RL Biochemistry 33:508-517(1994).

RN [7]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RX MEDLINE=95244421; PubMed=7727362;

RA Endrizzi J.A., Breddam K., Remington S.J.;

RT "2.8-A structure of yeast serine carboxypeptidase.";

RL Biochemistry 33:11106-11120(1994).

CC -!- FUNCTION: Involved in degradation of small peptides. Digests

CC preferentially peptides containing an aliphatic or hydrophobic

CC residue in P1' position, as well as methionine, leucine or

CC phenylalanine in P1 position of ester substrate.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with broad

CC specificity.

CC -!- ENZYME REGULATION: Inhibited by ZPCK.

CC -!- SUBCELLULAR LOCATION: Lysosome-like vacuoles.

CC -!- PTM: Enters the endoplasmic reticulum as an inactive zymogen and

CC is modified by four N-linked core oligosaccharides, giving rise to

CC a precursor known as P1 (67 kDa). As P1 transits through the

CC Golgi, extension of its core oligosaccharides leads to the Golgi-

CC modified P2 precursor (69 kDa). P2 is sorted away from secretory

CC proteins at or beyond a late Golgi compartment and is subsequently

CC delivered to the vacuole via a prevacuolar endosome-like

CC compartment. Upon arrival in the vacuole, the N-terminal

CC prosegment of P2 is cleaved to yield the enzymatically active

CC mature vacuolar form of CPY (61 kDa).

CC -!- SIMILARITY: Belongs to the peptidase S10 family.

CC -!- DATABASE: NAME=Worthington enzyme manual;

CC WWW="http://www.worthington-biochem.com/COY/".

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

DR EMBL; M15482; AAA34902.1; -; Genomic_DNA.

DR EMBL; X80836; CAA56806.1; -; Genomic_DNA.

DR PIR; A26597; CPBYY.

DR PDB; 1CPY; X-ray; @=112-532.

DR PDB; 1WPX; X-ray; A=112-532.

DR PDB; 1YSC; X-ray; @=112-532.

DR GermOnline; 142974; -.

DR MEROPS; S10.001; -.

DR Ensembl; YMR297W; Saccharomyces cerevisiae.

DR SGD; S000004912; PRCL.

DR GO; GO:0005783; C:endoplasmic reticulum; IDA.

DR GO; GO:000328; C:vacuolar lumen (sensu Fungi); TAS.

DR GO; GO:0004186; F:carboxypeptidase C activity; TAS.

DR GO; GO:0007039; P:vacuolar protein catabolism; TAS.

DR InterPro; IPR008442; Carbppep_N.

DR InterPro; IPR001563; Peptidase_S10.

DR InterPro; IPR000379; Ser_estrs.

DR PANTHER; PTHR11802; Peptidase_S10; 1.

DR Pfam; PF05388; Carbppep_Y_N; 1.

DR Pfam; PF00450; Peptidase_S10; 1.

DR PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; PD001189; Peptidase_S10; 1.

DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.

KW 3D-structure; Carboxypeptidase; Complete proteome;

KW Direct protein sequencing; Glycoprotein; Hydrolase; Signal;

KW Vacuole; Zymogen.

FT SIGNAL 1 20 Potential.

FT PROPEP 21 111

FT CHAIN 112 532 Carboxypeptidase Y.

FT MOTIF 24 27 Vacuolar targeting signal.

FT ACT_SITE 257 257

FT ACT_SITE 449 449 By similarity.

FT BINDING 508 508 Substrate.

FT BINDING 509 509 Substrate.

FT CARBOHYD 124 124 N-linked (GlcNAc. .).

FT CARBOHYD 198 198 N-linked (GlcNAc. .).

FT CARBOHYD 279 279 N-linked (GlcNAc. .).

FT CARBOHYD 479 479 N-linked (GlcNAc. .).

FT DISULFID 167 409

FT DISULFID 304 318

FT DISULFID 328 351

FT DISULFID 335 344

FT DISULFID 373 379

FT MUTAGEN 508 508

FT CONFLICT 260 261 H->A,R: Inactivates enzyme.

FT CONFLICT 389 389 GH -> HG (in Ref. 3).

FT CONFLICT 529 529 Y -> E (in Ref. 3).

FT STRAND 114 114 G -> D (in Ref. 3).

FT HELIX 116 118

FT STRAND 129 134

FT TURN 135 138

FT STRAND 139 146

FT TURN 152 154

FT STRAND 157 161

FT TURN 164 166

FT STRAND 167 167

FT TURN 169 169

FT HELIX 170 173

FT TURN 174 177

FT STRAND 180 183

FT TURN 184 186

FT STRAND 187 190

FT TURN 192 193

FT HELIX 195 198

FT STRAND 200 202

FT TURN 208 209

FT TURN 211 212

FT STRAND 214 215

FT HELIX 224 240

FT TURN 242 243

FT TURN 245 248

FT STRAND 251 256

FT TURN 257 258

FT HELIX 259 270

FT TURN 271 272

FT STRAND 282 286

FT HELIX 292 295

FT HELIX 296 298

FT HELIX 299 303

FT TURN 304 305

FT HELIX 315 338

FT HELIX 341 355

FT TURN 356 356

FT HELIX 357 362

FT STRAND 365 365

FT TURN 366 367

FT STRAND 368 368

FT TURN 381 381

FT HELIX 382 391

FT HELIX 393 398

FT TURN 399 400

FT STRAND 409 409

FT HELIX 411 418

FT TURN 419 421


```

FT HELIX 422 424
FT TURN 427 427
FT HELIX 428 436
FT TURN 437 438
FT STRAND 441 446
FT TURN 447 448
FT TURN 451 452
FT HELIX 454 463
FT TURN 467 468
FT HELIX 469 474
FT STRAND 478 481
FT TURN 483 485
FT STRAND 488 492
FT STRAND 494 495

Query Match 57.7%; Score 45; DB 1; Length 532;
Best Local Similarity 66.7%; Pred. NO. 90;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
Db 522 VNEWIHGGF 530

RESULT 21
Q7V2A4_PROMP PRELIMINARY; PRT; 154 AA.
AC Q7V2A4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PMM0579;
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572091; CAE19038.1; -; Genomic_DNA.
DR InterPro; IPR007115; 6 PTP_synth.
DR PIRSF; PIRSF006113; PTP_syn; 1.
DR ProDom; PD004049; PTPS_hypoth; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 154 AA; 18598 MW; D5E9B5EE909C89BC CRC64;

Query Match 56.4%; Score 44; DB 2; Length 154;
Best Local Similarity 50.0%; Pred. NO. 37;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 19 CSHRQWRHDGHC 30

RESULT 22
Q9ZPU0_ARATH
ID Q9ZPU0_ARATH PRELIMINARY; PRT; 171 AA.
AC Q9ZPU0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein At2g13980.
GN Name=At2g13980;
```

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006528; AAD19777.1; -; Genomic_DNA.
DR PIR; G84512; G84512.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR InterPro; IPR012337; RNaseH fold.
KW Hydrolase; Hypothetical protein; Nuclease.
SQ SEQUENCE 171 AA; 19347 MW; D615373DB0453A5A CRC64;

Query Match 56.4%; Score 44; DB 2; Length 171;
Best Local Similarity 58.3%; Pred. NO. 42;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 156 CTLPWLHQHYC 167

RESULT 23
Q6CBE1_YARLI
ID Q6CBE1_YARLI PRELIMINARY; PRT; 539 AA.
AC Q6CBE1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to CA0795|IPF15641 Candida albicans IPF15641 unknown
DE function.
GN OrderedLocusNames=YALI0C19682g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR382129; CAG82341.1; -; Genomic_DNA.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
KW Complete proteome.
SQ SEQUENCE 539 AA; 61646 MW; AF9764FD74AEB954 CRC64;
```

Query Match 56.4%; Score 44; DB 2; Length 539;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 376 CSLAEWFLGGF 386
|:| | | | | | |

RESULT 24
Q61MS4 CAEBR
ID Q61MS4 CAEBR PRELIMINARY; PRT; 731 AA.
AC Q61MS4;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG08359.
GN Name=CBG08359;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
family.
DR EMBL; CAAC01000035; CAE63810.1; -; Genomic_DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Hypothetical protein.
SQ SEQUENCE 731 AA; 81445 MW; C071504A05B95631 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 731;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
Db 517 LREWLEGGY 525
| | | | | | |

RESULT 25
Q628B7 CAEBR
ID Q628B7 CAEBR PRELIMINARY; PRT; 1540 AA.
AC Q628B7;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG00449.
GN Name=CBG00449;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC01000004; CAE57481.1; -; Genomic_DNA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 3.
KW EGF-like domain; Hypothetical protein.
SQ SEQUENCE 1540 AA; 168546 MW; 48CECC95069F6E60 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 1540;
Best Local Similarity 63.6%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
Db 35 CTIRECLNGGY 45
|:| | | | | | |

RESULT 26
Q61EJ2 CAEBR
ID Q61EJ2 CAEBR PRELIMINARY; PRT; 4856 AA.
AC Q61EJ2;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG12050.
GN Name=CBG12050;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC01000059; CAE66705.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF00057; Ldl_recept_a; 28.
DR Pfam; PF00058; Ldl_recept_b; 18.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 25.
DR SMART; SM00179; EGF_CA; 8.
DR SMART; SM00192; LDLa; 32.
DR SMART; SM00135; LY; 27.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS50026; EGF_3; 9.
DR PROSITE; PS01209; LDLRA_1; 18.
DR PROSITE; PS50068; LDLRA_2; 24.
KW Hypothetical protein.
SQ SEQUENCE 4856 AA; 541409 MW; D10DA26DC9A5CC8A CRC64;

Query Match 56.4%; Score 44; DB 2; Length 4856;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
Db 4464 CKDRECLNGGFC 4475
| | | | | | |

RESULT 27

Q7U636_SYNPX
ID Q7U636_SYNPX PRELIMINARY; PRT; 158 AA.
AC Q7U636;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SYNW1504;
OS Synecococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=84588;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.,
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569693; CAE08019.1; -; Genomic_DNA.
DR InterPro; IPR007115; 6 PTP_synth.
DR PIRSF; PIRSF006113; PTP_syn; 1.
DR ProDom; PD004049; PTPS_hypoth; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 158 AA; 18232 MW; DB0A4EDAD8831AAB CRC64;

Query Match 55.1%; Score 43; DB 2; Length 158;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | |
DB 19 CCHRQMRHAGHC 30

RESULT 28
Q9B598_CHILA
ID Q9B598_CHILA PRELIMINARY; PRT; 182 AA.
AC Q9B598;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Chinchilla lanigera (Chinchilla).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
OC Hystriognathi; Chinchillidae; Chinchilla.
OX NCBI_TaxID=34839;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skin;
RA Spottorno A.E., Valladares J.P., Zuleta C., Marin J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF244380; AAK28364.1; -; Genomic_DNA.
DR SMR; Q9B598; 13-182.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006910; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.

DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 182
SQ SEQUENCE 182 AA; 20594 MW; CAA7A70EFF51374E CRC64;

Query Match 55.1%; Score 43; DB 2; Length 182;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
| | | | |
DB 163 TLEEWIWWGF 172

RESULT 29
Q5V5F2_HALMA
ID Q5V5F2_HALMA PRELIMINARY; PRT; 328 AA.
AC Q5V5F2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative transporter.
GN OrderedLocusNames=rrnAC0187;
OS Haloarcula marismortui (Haloacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=2238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43049;
RX PubMed=15520287; DOI=10.1101/gr.2700304;
RA Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G.,
RA Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P.,
RA Date S.V., Marcotte E., Hood L., Ng W.V.;
RT "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
RT the Dead Sea.";
RL Genome Res. 14:2221-2234(2004).
DR EMBL; AY596297; AAV45250.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 328 AA; 35627 MW; 4DC488FB91107046 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 328;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
| | | | |
DB 293 TLEWIWWGW 302

RESULT 30
Q9B640_9RODE
ID Q9B640_9RODE PRELIMINARY; PRT; 380 AA.
AC Q9B640;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=Cytb;
OS Eothenomys andersoni.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Cricetidae; Arvicolinae; Eothenomys.
OX NCBI_TaxID=82466;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Iwasa M.A., Suzuki H.;
RT "Evolutionary networks of maternal and paternal gene lineages in voles
RT (Eothenomys) endemic to Japan.";

RL J. Mammal. 83:852-865(2002).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-cl complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AB037303; BAB40558.1; -; Genomic_DNA.
DR SMR; Q9B640; 2-377.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 380
SQ SEQUENCE 380 AA; 42882 MW; 188E774B649B44BD CRC64;

Query Match 55.1%; Score 43; DB 2; Length 380;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
Db 159 TLEEWIWGGF 168

RESULT 31
Q8WF54_9BIVA PRELIMINARY; PRT; 440 AA.
ID Q8WF54_9BIVA PRELIMINARY; PRT; 440 AA.
AC Q8WF54;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b.
GN Name=cytb;
OS Venerupis (Ruditapes) philippinarum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Venerupis; Ruditapes.
OX NCBI_TaxID=129788;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Gonad;
RA Okazaki M., Ueshima R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-cl complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AB065374; BAB83775.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
SQ SEQUENCE 440 AA; 49604 MW; 93C3A76788C8E2FD CRC64;

Query Match 55.1%; Score 43; DB 2; Length 440;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 TLREWLHGGFC 12
Db 170 TLMIEWVGGHC 180

RESULT 32
Q4ICS6_GIBZE
ID Q4ICS6_GIBZE PRELIMINARY; PRT; 558 AA.
AC Q4ICS6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG04982.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000200; EAA74849.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 558 AA; 61297 MW; FACB1CEFF89AD96 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 558;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLREWLHGG 10

Db 107 SMRNWLHGG 115
:| | | | |
Query Match 55.1%; Score 43; DB 2; Length 583;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 5 EWLHGGFC 12
:| | | | |
Db 459 QWLHGGRC 466
RESULT 34
Q59GS6 HUMAN
ID Q59GS6 HUMAN PRELIMINARY; PRT; 605 AA.
AC Q59GS6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Diacylglycerol kinase, gamma 90kDa variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209033; BAD92270.1; -; mRNA.
KW Kinase.
FT NON TER 1 1
SQ SEQUENCE 605 AA; 68332 MW; 18E1846CA4366F2E CRC64;
Query Match 55.1%; Score 43; DB 2; Length 605;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 5 EWLHGGFC 12
:| | | | |
Db 459 QWLHGGRC 466

Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 2 TLREWLHGG 10
:| | | | |
Db 244 SLOEWHVG 252
RESULT 35
Q5FWG1 HUMAN
ID Q5FWG1 HUMAN PRELIMINARY; PRT; 766 AA.
AC Q5FWG1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DGKG protein.
GN Name=DGKG;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymph;
RA Director MGC Project;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC089411; AAH89411.1; -; mRNA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000756; DAGKa.
DR InterPro; IPR001206; DAGKc.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR Pfam; PF00130; Cl_1; 2.
DR Pfam; PF00609; DAGK_acc; 1.
DR Pfam; PF00781; DAGK_cat; 1.
DR Pfam; PF00036; efhand; 2.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR ProDom; PD002939; DAGKa; 1.
DR ProDom; PD005043; DAGKc; 1.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00109; Cl; 2.
DR SMART; SM00045; DAGKa; 1.
DR SMART; SM00046; DAGKc; 1.
DR SMART; SM00054; EFh; 2.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.

```
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
KW Calcium; Repeat.
SQ SEQUENCE 766 AA; 86303 MW; 1D3F8C78BD59219F CRC64;

Query Match 55.1%; Score 43; DB 2; Length 766;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLREWLG 10
Db 241 SLQEWVHGG 249

RESULT 36
DGKG HUMAN
ID DGKG HUMAN STANDARD; PRT; 791 AA.
AC P49619;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Diacylglycerol kinase, gamma (EC 2.7.1.107) (Diglyceride kinase) (DGK-
DE gamma) (DAG kinase gamma).
DE Name=DGKG; Synonyms=DAGK3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94308084; PubMed=8034597;
RA Kai M., Sakane F., Imai S.-I., Wada I., Kanoh H.;
RT "Molecular cloning of a diacylglycerol kinase isozyme predominantly
RT expressed in human retina with a truncated and inactive enzyme
RT expression in most other human cells.";
RL J. Biol. Chem. 269:18492-18498 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99168758; PubMed=10071200; DOI=10.1007/s004390050917;
RA Stoehr H., Klein J., Gehrig A., Koehler M.R., Jurkles B., Kellner U.,
RA Leo-Kottler B., Schmid M., Weber B.H.P.;
RT "Mapping and genomic characterization of the gene encoding
RT diacylglycerol kinase gamma (DAGK3): assessment of its role in
RT dominant optic atrophy (OPAL).";
RL Hum. Genet. 104:99-105 (1999).
CC -!- FUNCTION: Reverses the normal flow of glycerolipid biosynthesis by
CC phosphorylating diacylglycerol back to phosphatidic acid.
CC -!- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacyl-
CC sn-glycerol 3-phosphate.
CC -!- ENZYME REGULATION: Requires phosphatidylserine for maximal
CC activity.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Can be loosely bound to the
CC membranes.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P49619-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P49619-2; Sequence=VSP_001267;
CC Note=May be inactive;
CC -!- TISSUE SPECIFICITY: Predominantly expressed in retina and in a
CC much lesser extent in the brain. Other tissues contain extremely
CC low levels of DGK-gamma.
CC -!- SIMILARITY: Belongs to the eukaryotic diacylglycerol kinase
CC family.
CC -!- SIMILARITY: Contains 2 EF-hand domains.
CC -!- SIMILARITY: Contains 2 phorbol-ester/DAG-type zinc fingers.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; D26135; BAA05132.1; -; mRNA.
DR EMBL; AF020945; AAC04686.1; -; Genomic DNA.
DR EMBL; AF020922; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020923; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020924; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020925; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020926; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020927; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020928; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020929; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020930; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020931; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020932; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020933; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020934; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020935; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020936; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020937; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020938; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020939; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020940; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020941; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020942; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020943; AAC04686.1; JOINED; Genomic DNA.
DR EMBL; AF020944; AAC04686.1; JOINED; Genomic DNA.
DR PIR; A53691; A53691.
DR Ensembl; ENSG00000058866; Homo sapiens.
DR HGNC; HGNC:2853; DGKG.
DR MIM; 601854; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002219; DAG_PE_bd.
DR InterPro; IPR000756; DAGKa.
DR InterPro; IPR001206; DAGKc.
DR InterPro; IPR011992; EF-Hand type.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR Pfam; PF00130; Cl_1; 2.
DR Pfam; PF00609; DAGK_acc; 1.
DR Pfam; PF00781; DAGK_cat; 1.
DR Pfam; PF00036; ehand; 2.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR ProDom; PD002939; DAGKa; 1.
DR ProDom; PD005043; DAGKc; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND_1; 2.
DR PROSITE; PS0222; EF_HAND_2; 2.
DR PROSITE; PS00479; ZF_DAG_PE_1; 2.
DR PROSITE; PS00081; ZF_DAG_PE_2; 2.
KW Alternative splicing; Calcium; Kinase; Metal-binding;
KW Multigene family; Phorbol-ester binding; Polymorphism; Repeat;
KW Transferase; Zinc; Zinc-finger.
FT DOMAIN 175 210 EF-hand 1.
FT DOMAIN 220 255 EF-hand 2.
FT CA_BIND 188 199 1 (Potential).
FT CA_BIND 233 244 2 (Potential).
FT ZN_FING 271 321 Phorbol-ester/DAG-type 1.
FT ZN_FING 336 385 Phorbol-ester/DAG-type 2.
FT REGION 432 558 Catalytic-A (Potential).
FT REGION 578 752 Catalytic-B (Potential).
FT COMPBIAS 151 156 Poly-Ser.
FT VARSPLIC 451 475 Missing (in isoform Short).
FT VARIANT 142 142 /FTId=VSP_001267.
FT VARIANT 370 370 T -> S (in dbSNP:1004588).
FT VARIANT 370 370 R -> W (in dbSNP:3213770).
SQ SEQUENCE 791 AA; 88997 MW; C7DD07F5B285FF62 CRC64;

Query Match 55.1%; Score 43; DB 1; Length 791;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
```


RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK020589; BAB32140.1; -; mRNA.
DR EMBL; AK047294; BAC33018.1; -; mRNA.
DR Ensembl; ENSMUSG00000047363; Mus musculus.
DR MGI; MGI:1925867; Cstsd.
DR GO; GO:0005741; C:mitochondrial outer membrane; IDA.
DR GO; GO:0007006; P:mitochondrial membrane organization and bio. .; IDA.
KW Hypothetical protein.
SQ SEQUENCE 104 AA; 11683 MW; 17CC14DE2664C0C7 CRC64;

Query Match 54.5%; Score 42.5; DB 2; Length 104;
Best Local Similarity 52.9%; Pred. No. 44;
Matches 9; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

QY 1 CTLREW-----LHGFC 12
Db ||||| | |||||
36 CTLRSWGVARVLPGEFC 52

RESULT 39
Q9T204 BPLC2
ID Q9T204 BPLC2 PRELIMINARY; PRT; 58 AA.
AC Q9T204;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Minor structural protein (Fragment).
GN Name=l16;
OS Lactococcus bacteriophage c2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC c2-like viruses.
OX NCBI_TaxID=31537;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kotsonis S.E., Powell I.B., Lubbers M.W., Hamilton R.C., Jarvis A.W.,
RA Hillier A.J., Limsowtin G.K.Y., Davidson B.E.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AFL18396; AAD20609.1; -; Genomic_DNA.
FT NON_TER 1 1

SQ SEQUENCE 58 AA; 6881 MW; B7CA331E09F7F1B4 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 58;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 EWLHGGFC 12
Db :||:|:|
39 KWLGGYC 46

RESULT 40
Q85A85 CORMC
ID Q85A85 CORMC PRELIMINARY; PRT; 112 AA.
AC Q85A85;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=Cytb;
OS Corvus macrorhynchos (Jungle crow).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;
OC Corvinae; Corvus.
OX NCBI_TaxID=36249;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Iwasa M.A., Kryukov A.P., Kakizawa R., Suzuki H.;
RT "Local Differentiation of Mitochondrial Gene of Jungle Crow, Corvus macrorhynchos (Corvidae), in East and South Asia.";
RL Yamashina Choru Kenkyujo Kenkyu Hokoku 0:0-0(2002).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AB092441; BAC54970.1; -; Genomic DNA.
DR EMBL; AB092442; BAC54971.1; -; Genomic DNA.
DR EMBL; AB092464; BAC54993.1; -; Genomic DNA.
DR SMR; Q85A85; 1-112.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12480 MW; DDF3BF6BCA2FE347 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 112;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
Db ||||| |||||
46 TLVEWLWGGF 55

RESULT 41
Q8BG00 MOUSE
ID Q8BG00 MOUSE PRELIMINARY; PRT; 121 AA.
AC Q8BG00;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Mus musculus adult male aorta and vein cDNA, RIKEN full-length
DE enriched library, clone:A530088E08 product:hypothetical protein, full
DE insert sequence (Mus musculus 4 days neonate male adipose cDNA, RIKEN
DE full-length enriched library, clone:B430004P05 product:hypothetical
DE protein, full insert sequence).
GN Name=A430107D22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK041179; BAC30852.1; -; mRNA.
DR EMBL; AK046549; BAC32781.1; -; mRNA.
DR Ensembl; ENSMUSG00000053467; Mus musculus.
DR MGI; MGI:2444128; A430107D22Rik.
KW Hypothetical protein.
SQ SEQUENCE 121 AA; 13068 MW; DA6405F40B706FEE CRC64;

Query Match 53.8%; Score 42; DB 2; Length 121;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLREWLHGG 10
Db ||:|:|
47 TLQRWVHGG 55

RESULT 42
P92837_9HYME
ID P92837_9HYME PRELIMINARY; PRT; 128 AA.
AC P92837;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Ephedrus persicae.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonoidea;
OC Braconidae; Aphidinae; Ephedrus.

OX NCBI_TaxID=55917;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97330629; PubMed=9187088; DOI=10.1006/mpev.1996.0400;
RA Belshaw R.D., Quicke D.L.J.;
RT "A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae).";
RL Mol. Phylogenet. Evol. 7:281-293(1997).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; Z83630; CAB05972.1; -; Genomic DNA.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0016020; C: membrane; IEA.
DR GO: 0005746; C: mitochondrial electron transport chain; IEA.
DR GO: 0005739; C: mitochondrial electron transport chain; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 15261 MW; 62009B9B3D42E3C5 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 128;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
Db 12 TLVEWLWGGF 21

RESULT 43
Q7YAZ1_9SAUR
ID Q7YAZ1_9SAUR PRELIMINARY; PRT; 134 AA.
AC Q7YAZ1;
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Colobosaura modesta.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Teiioidea;
OC Gymnophthalmidae; Colobosaura.
OX NCBI_TaxID=174786;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22980429; PubMed=14615195; DOI=10.1016/S1055-7903(03)00142-8;
RA Whiting A.S., Bauer A.M., Sites J.W. Jr.;
RT "Phylogenetic relationships and limb loss in sub-Saharan African
RT scincine lizards (Squamata: Scincidae).";
RL Mol. Phylogenet. Evol. 29:582-598(2003).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,

cytochrome c1 and the Rieske protein (By similarity).
-!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AY217799; AAP45101.1; -; Genomic DNA.
DR SMR; Q7YAZ1; 1-134.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0016020; C: membrane; IEA.
DR GO: 0005746; C: mitochondrial electron transport chain; IEA.
DR GO: 0005739; C: mitochondrial electron transport chain; IEA.
DR GO: 0016491; F: oxidoreductase activity; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15106 MW; 97C12F6552055EE8 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 134;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
Db 23 TLVEWLWGGF 32

RESULT 44
Q9MMC3_GUIGU
ID Q9MMC3_GUIGU PRELIMINARY; PRT; 144 AA.
AC Q9MMC3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Guira guira (Guira cuckoo).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Cuculiformes; Crotophagidae; Guira.
OX NCBI_TaxID=30392;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20179528; PubMed=10712848; DOI=10.1006/mpev.1999.0715;
RA Johnson K.P., Goodman S.M., Lanyon S.M.;
RT "A phylogenetic study of the malagasy couas with insights into cuckoo
RT relationships.";
RL Mol. Phylogenet. Evol. 14:436-444(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF204998; AAF6885.1; -; Genomic DNA.
DR SMR; Q9MMC3; 1-144.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0016020; C: membrane; IEA.
DR GO: 0005746; C: mitochondrial electron transport chain; IEA.
DR GO: 0005739; C: mitochondrial electron transport chain; IEA.
DR GO: 0016491; F: oxidoreductase activity; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR GO: 0006118; P: electron transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.

```
FT  NON_TER      1      1
FT  NON_TER     144     144
SQ  SEQUENCE    144 AA; 16103 MW; 2FCE9A0B86DD39C9 CRC64;

Query Match      53.8%; Score 42; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TLREWHLGGF 11
      |||||
Db      127 TLVEWLWGGF 136

RESULT 45
Q9MMC4 9AVES
ID  Q9MMC4 9AVES PRELIMINARY; PRT; 144 AA.
AC  Q9MMC4;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Crotophaga major.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Cuculiformes; Crotophagidae;
OC  Crotophaga.
OX  NCBI_TaxID=48617;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=20179528; PubMed=10712848; DOI=10.1006/mpev.1999.0715;
RA  Johnson K.P.; Goodman S.M.; Lanyon S.M.;
RT  "A phylogenetic study of the malagasy couas with insights into cuckoo
RL  Mol. Phylogenet. Evol. 14:436-444(2000).
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-c1 complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AF204997; AAF68884.1; -; Genomic_DNA.
DR  SMR; Q9MMC4; 1-144.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016020; C:membrane; IEA.
DR  GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR  GO; GO:0005739; C:mitochondrion; IEA.
DR  GO; GO:0016491; F:oxidoreductase activity; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR005797; Cytb_b6_N.
DR  Pfam; PF00033; Cytochrom_b_N; 1.
DR  PROSITE; PS51002; CYTB_NTER; 1.
KW  Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW  Respiratory chain; Transmembrane; Transport.
FT  NON_TER      1      1
FT  NON_TER     144     144
SQ  SEQUENCE    144 AA; 16111 MW; 5423190720CCFFA4 CRC64;

Query Match      53.8%; Score 42; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TLREWHLGGF 11
      |||||
Db      127 TLVEWLWGGF 136

RESULT 46
Q4VTU6_9SAUR
ID  Q4VTU6_9SAUR PRELIMINARY; PRT; 193 AA.
AC  Q4VTU6;
```

```
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT  13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Gymnodactylus darwini.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
OC  Gymnodactylus.
OX  NCBI_TaxID=298113;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=MRTA2;
RA  Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
RA  Yassuda Y.Y., Sites J.W. Jr.;
RT  "Phylogeography and species limits in the Gymnodactylus darwini
RT  complex (Gekkonidae, Squamata): genetic structure coincides with river
RT  systems in the Brazilian Atlantic Forest.";
RL  Biol. J. Linn. Soc. Lond. 85:13-26(2005).
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-c1 complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AY630394; AAV54568.1; -; Genomic_DNA.
DR  InterPro; IPR005797; Cytb_b6_N.
DR  Pfam; PF00033; Cytochrom_b_N; 1.
DR  PROSITE; PS51002; CYTB_NTER; 1.
KW  Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW  Respiratory chain; Transmembrane; Transport.
FT  NON_TER      1      1
FT  NON_TER     193     193
SQ  SEQUENCE    193 AA; 21639 MW; CEAA98C5C321D19C CRC64;

Query Match      53.8%; Score 42; DB 2; Length 193;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TLREWHLGGF 11
      |||||
Db      141 TLVEWLWGGF 150

RESULT 47
Q7YB81_9SYLV
ID  Q7YB81_9SYLV PRELIMINARY; PRT; 196 AA.
AC  Q7YB81;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Alcippe morrisonia.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; Alcippe.
OX  NCBI_TaxID=204791;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Pasquet E., Cibois A., Kalyakin M.V., Bourdon E.;
RL  Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-c1 complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AY128557; AAN01293.1; -; Genomic_DNA.
DR  SMR; Q7YB81; 1-196.
```

DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 196 196
SQ SEQUENCE 196 AA; 21764 MW; A8476B4C2F34447E8 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 196;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||| |||
Db 22 TLVEWLWGGF 31

RESULT 48
Q4VTU5_9SAUR
ID Q4VTU5_9SAUR PRELIMINARY; PRT; 203 AA.
AC Q4VTU5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cytochrome b (Fragment).
OS Gymnodactylus darwini.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
OC Gymnodactylus.
OX NCBI_TaxID=298113;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MRTA3;
RA Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
RA Yassuda Y.Y., Sites J.W. Jr.;
RT "Phylogeography and species limits in the Gymnodactylus darwini
RT complex (Gekkonidae, Squamata): genetic structure coincides with river
RT systems in the Brazilian Atlantic Forest.";
RL Biol. J. Linn. Soc. Lond. 85:13-26(2005).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AY630395; AAV54569.1; -; Genomic_DNA.
DR EMBL; AY630393; AAV54567.1; -; Genomic_DNA.
DR InterPro; IPR005797; Cytb_b6_N.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 203 203
SQ SEQUENCE 203 AA; 22852 MW; F4FC8230EC586FCA CRC64;

Query Match 53.8%; Score 42; DB 2; Length 203;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||| |||
Db 136 TLVEWLWGGF 145
RESULT 49
Q4VTU4_9SAUR
ID Q4VTU4_9SAUR PRELIMINARY; PRT; 213 AA.
AC Q4VTU4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cytochrome b (Fragment).
OS Gymnodactylus darwini.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
OC Gymnodactylus.
OX NCBI_TaxID=298113;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MRTA5, and MRTA1;
RA Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
RA Yassuda Y.Y., Sites J.W. Jr.;
RT "Phylogeography and species limits in the Gymnodactylus darwini
RT complex (Gekkonidae, Squamata): genetic structure coincides with river
RT systems in the Brazilian Atlantic Forest.";
RL Biol. J. Linn. Soc. Lond. 85:13-26(2005).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AY630396; AAV54570.1; -; Genomic_DNA.
DR EMBL; AY630393; AAV54567.1; -; Genomic_DNA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 213 213
SQ SEQUENCE 213 AA; 23895 MW; 2B6EA26AF385F7DD CRC64;

Query Match 53.8%; Score 42; DB 2; Length 213;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
||| ||| |||
Db 136 TLVEWLWGGF 145

RESULT 50
Q34157_CORMC
ID Q34157_CORMC PRELIMINARY; PRT; 215 AA.
AC Q34157;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Corvus macrorhynchos (Jungle crow).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;
OC Corvinae; Corvus.
OX NCBI_TaxID=36249;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Muscle;
RA Chikuni K., Minaka N., Ikenaga H.;
RT "Molecular phylogeny of some Passeriformes, based on cytochrome b
RT sequences.";
RL Yamashina Chorui Kenkyujo Kenkyu Hokoku 28:1-8(1995).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; D38313; BAA07429.1; -; Genomic_DNA.
DR SMR; Q34157; 1-215.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 215 215
SQ SEQUENCE 215 AA; 24057 MW; 65D3C67C2F496D1D CRC64;

Query Match 53.8%; Score 42; DB 2; Length 215;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||
Db 107 TLVEWLWGGF 116

Search completed: May 12, 2006, 10:49:09
Job time : 99.9231 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 101.709 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-53
Perfect score: 81
Sequence: 1 CSLQEFSLSHGGYVC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	14	2	AAW09535 Thrombopo
2	81	100.0	14	2	AAW36686 Thrombopo
3	81	100.0	14	3	AAB16997 TPO-mimet
4	81	100.0	14	4	AAU25905 Human thr
5	81	100.0	14	5	ABB72883 TPO mimet
6	81	100.0	14	7	ADJ73034 TPO mimet
7	81	100.0	14	8	ADJ52669 CH1 delet
8	81	100.0	14	8	ADJ51630 CH1 delet
9	49	60.5	453	8	ADT60670 Plant pol
10	48	59.3	344	3	AAY96482 Human top
11	48	59.3	660	2	AAW81503 Dead Box
12	48	59.3	660	6	ABO07237 Human p53
13	48	59.3	660	8	ADL13058 Human ste
14	48	59.3	660	8	ADM33393 Human PRO
15	48	59.3	660	9	ADW08765 Human PRO
16	48	59.3	660	9	AEA24054 Human PRO
17	48	59.3	662	2	AAW81501 Dead Box
18	48	59.3	662	2	AAW81502 Dead Box
19	48	59.3	662	3	AAY96483 Human DBX
20	48	59.3	662	5	ABB57045 Mouse isc
21	48	59.3	662	6	ABO07236 Human p53
22	48	59.3	662	6	ABO52958 Human spl
23	48	59.3	662	8	ADN04580 Antipsori
24	48	59.3	662	8	ADN03714 Antipsori

98	44	54.3	490	6	ADB14798	Adb14798 Human PRO	171	44	54.3	490	7	ADC34278	Adc34278 Human sec
99	44	54.3	490	6	ADB18759	Adb18759 Novel hum	172	44	54.3	490	7	ADC29333	Adc29333 Human sec
100	44	54.3	490	6	ADA93974	Ada93974 Human PRO	173	44	54.3	490	7	ADC28864	Adc28864 Human sec
101	44	54.3	490	6	ADB19870	Adb19870 Novel hum	174	44	54.3	490	7	ADC40749	Adc40749 Human sec
102	44	54.3	490	6	ADB13182	Adb13182 Human PRO	175	44	54.3	490	7	ADC19406	Adc19406 Human sec
103	44	54.3	490	6	ABO43303	Abo43303 Novel hum	176	44	54.3	490	7	ADC33854	Adc33854 Human sec
104	44	54.3	490	6	ADA74436	Ada74436 Human PRO	177	44	54.3	490	7	ADC12924	Adc12924 Human sec
105	44	54.3	490	6	ADA42313	Ada42313 Human sec	178	44	54.3	490	7	ADC50351	Adc50351 Novel hum
106	44	54.3	490	6	ADB24669	Adb24669 Human PRO	179	44	54.3	490	7	ADC71898	Adc71898 Novel hum
107	44	54.3	490	6	ADA82193	Ada82193 Human PRO	180	44	54.3	490	7	ADC59877	Adc59877 Novel hum
108	44	54.3	490	6	ADA75156	Ada75156 Human PRO	181	44	54.3	490	7	ADC52884	Adc52884 Novel hum
109	44	54.3	490	6	ADA85234	Ada85234 Novel hum	182	44	54.3	490	7	ADC57238	Adc57238 Novel hum
110	44	54.3	490	6	ADA84682	Ada84682 Novel hum	183	44	54.3	490	7	ADC60429	Adc60429 Novel hum
111	44	54.3	490	6	ABO17516	Abo17516 Human PRO	184	44	54.3	490	7	ADC50904	Adc50904 Novel hum
112	44	54.3	490	6	ADB29938	Adb29938 Human PRO	185	44	54.3	490	7	ADC65431	Adc65431 Human PRO
113	44	54.3	490	6	ADA80466	Ada80466 Human PRO	186	44	54.3	490	7	ADC54529	Adc54529 Novel hum
114	44	54.3	490	6	ADA75708	Ada75708 Human PRO	187	44	54.3	490	7	ADC53490	Adc53490 Novel hum
115	44	54.3	490	6	ADA46933	Ada46933 Human PRO	188	44	54.3	490	7	ADC59013	Adc59013 Novel hum
116	44	54.3	490	6	ADB25229	Adb25229 Human PRO	189	44	54.3	490	7	ADC55891	Adc55891 Novel hum
117	44	54.3	490	6	ADA93405	Ada93405 Human PRO	190	44	54.3	490	7	ADC58461	Adc58461 Novel hum
118	44	54.3	490	6	ADB26755	Adb26755 Human PRO	191	44	54.3	490	7	ADC12376	Adc12376 Human sec
119	44	54.3	490	6	ADB31042	Adb31042 Human PRO	192	44	54.3	490	7	ADD03135	Adc03135 Novel hum
120	44	54.3	490	6	ADA60970	Ada60970 Homo sapi	193	44	54.3	490	7	ADC90127	Adc90127 Novel hum
121	44	54.3	490	6	ADB24117	Adb24117 Human PRO	194	44	54.3	490	7	ADC69546	Adc69546 Human PRO
122	44	54.3	490	6	ADA96446	Ada96446 Human PRO	195	44	54.3	490	7	ADC48435	Adc48435 Human PRO
123	44	54.3	490	6	ADA81018	Ada81018 Human PRO	196	44	54.3	490	7	ADD09964	Adc09964 Human PRO
124	44	54.3	490	6	ADA95894	Ada95894 Human PRO	197	44	54.3	490	7	ADD04539	Adc04539 Novel hum
125	44	54.3	490	6	ADB26203	Adb26203 Human PRO	198	44	54.3	490	7	ADC80495	Adc80495 Novel hum
126	44	54.3	490	6	ADB21688	Adb21688 Novel hum	199	44	54.3	490	7	ADD11002	Adc11002 Human PRO
127	44	54.3	490	7	ADA77467	Ada77467 Human PRO	200	44	54.3	490	7	ADC47883	Adc47883 Human PRO
128	44	54.3	490	7	ADB18207	Adb18207 Human PRO	201	44	54.3	490	7	ADD04931	Adc04931 Human sec
129	44	54.3	490	7	ADA86890	Ada86890 Novel hum	202	44	54.3	490	7	ADC79943	Adc79943 Novel hum
130	44	54.3	490	7	ADA16592	Ada16592 Human sec	203	44	54.3	490	7	ADD09412	Adc09412 Human PRO
131	44	54.3	490	7	ADA13021	Ada13021 Human sec	204	44	54.3	490	7	ADD03937	Adc03937 Human sec
132	44	54.3	490	7	ADA41889	Ada41889 Human sec	205	44	54.3	490	7	ADD03513	Adc03513 Human sec
133	44	54.3	490	7	ADA87993	Ada87993 Novel hum	206	44	54.3	490	7	ADD41125	Adc41125 Novel hum
134	44	54.3	490	7	ADA46381	Ada46381 Novel hum	207	44	54.3	490	7	ADD52264	Adc52264 Human PRO
135	44	54.3	490	7	ADA17236	Ada17236 Human sec	208	44	54.3	490	7	ADD53004	Adc53004 Human PRO
136	44	54.3	490	7	ADA42739	Ada42739 Human sec	209	44	54.3	490	7	ADD53556	Adc53556 Novel hum
137	44	54.3	490	7	ADB28411	Adb28411 Human PRO	210	44	54.3	490	7	ADD51712	Adc51712 Human PRO
138	44	54.3	490	7	ADB28963	Adb28963 Human PRO	211	44	54.3	490	7	ADD02511	Adc02511 Human PRO
139	44	54.3	490	7	ADA76915	Ada76915 Human PRO	212	44	54.3	490	7	ADD01945	Adc01945 Human PRO
140	44	54.3	490	7	ADA88545	Ada88545 Novel hum	213	44	54.3	490	7	ADD54127	Adc54127 Novel hum
141	44	54.3	490	7	ADA97550	Ada97550 Human PRO	214	44	54.3	490	7	ADD92444	Adc92444 Human PRO
142	44	54.3	490	7	ADB27307	Adb27307 Human PRO	215	44	54.3	490	7	ADD91340	Adc91340 Human PRO
143	44	54.3	490	7	ADB22240	Adb22240 Novel hum	216	44	54.3	490	7	ADE03954	Ade03954 Human PRO
144	44	54.3	490	7	ABO17577	Abo17577 Human PRO	217	44	54.3	490	7	ADE32251	Ade32251 Novel hum
145	44	54.3	490	7	ADA66931	Ada66931 Human PRO	218	44	54.3	490	7	ADE22183	Ade22183 Human PRO
146	44	54.3	490	7	ADB22792	Adb22792 Human PRO	219	44	54.3	490	7	ADD79407	Adc79407 Human PRO
147	44	54.3	490	7	ADB23565	Adb23565 Human PRO	220	44	54.3	490	7	ADE41943	Ade41943 Human PRO
148	44	54.3	490	7	ADA92287	Ada92287 Novel hum	221	44	54.3	490	7	ADE17760	Ade17760 Human PRO
149	44	54.3	490	7	ADB15350	Adb15350 Human PRO	222	44	54.3	490	7	ADD91892	Adc91892 Human PRO
150	44	54.3	490	7	ADB38602	Adb38602 Novel hum	223	44	54.3	490	7	ADE33355	Ade33355 Novel hum
151	44	54.3	490	7	ADB38050	Adb38050 Novel hum	224	44	54.3	490	7	ADD79959	Adc79959 Human PRO
152	44	54.3	490	7	ADB66522	Adb66522 Novel hum	225	44	54.3	490	7	ADD79959	Adc79959 Human PRO
153	44	54.3	490	7	ADB89602	Adb89602 Human PRO	226	44	54.3	490	7	ADD92996	Adc92996 Human PRO
154	44	54.3	490	7	ADB90334	Adb90334 Human PRO	227	44	54.3	490	7	ADE19416	Ade19416 Human PRO
155	44	54.3	490	7	ADB77658	Adb77658 Human sec	228	44	54.3	490	7	ADE34765	Ade34765 Human sec
156	44	54.3	490	7	ADB39435	Adb39435 Novel hum	229	44	54.3	490	7	ADE18864	Ade18864 Human PRO
157	44	54.3	490	7	ADB74794	Adb74794 Human sec	230	44	54.3	490	7	ADE43060	Ade43060 Human PRO
158	44	54.3	490	7	ADB47058	Adb47058 Novel hum	231	44	54.3	490	7	ADD95849	Adc95849 Human PRO
159	44	54.3	490	7	ADB86665	Adb86665 Human PRO	232	44	54.3	490	7	ADE22735	Ade22735 Human PRO
160	44	54.3	490	7	ADB77270	Adb77270 Novel hum	233	44	54.3	490	7	ADD78853	Adc78853 Human PRO
161	44	54.3	490	7	ADB34427	Adb34427 Human PRO	234	44	54.3	490	7	ADE32803	Ade32803 Novel hum
162	44	54.3	490	7	ADB35531	Adb35531 Human PRO	235	44	54.3	490	7	ADE42495	Ade42495 Human PRO
163	44	54.3	490	7	ADB33875	Adb33875 Human PRO	236	44	54.3	490	7	ADD80511	Adc80511 Human PRO
164	44	54.3	490	7	ADB34979	Adb34979 Human PRO	237	44	54.3	490	7	ADD89539	Adc89539 Human PRO
165	44	54.3	490	7	ADB36083	Adb36083 Human PRO	238	44	54.3	490	7	ADE40823	Ade40823 Human PRO
166	44	54.3	490	7	ADB46478	Adb46478 Novel hum	239	44	54.3	490	7	ADE04622	Ade04622 Human PRO
167	44	54.3	490	7	ADC28440	Adc28440 Human sec	240	44	54.3	490	7	ADE92751	Ade92751 Human PRO
168	44	54.3	490	7	ADC39640	Adc39640 Human sec	241	44	54.3	490	7	ADG21460	Adg21460 Novel hum
169	44	54.3	490	7	ADC40154	Adc40154 Human sec	242	44	54.3	490	7	ADG23101	Adg23101 Novel hum
170	44	54.3	490	7	ADC18982	Adc18982 Human sec	243	44	54.3	490	7	ADF97436	Adf97436 Human PRO

244	44	54.3	490	7	ADG80500	Adg80500	Human	PRO	317	44	54.3	490	8	ADG82156	Adg82156	Human	PRO
245	44	54.3	490	7	ADG79948	Adg79948	Human	PRO	318	44	54.3	490	8	ADG57395	Adg57395	Novel	hum
246	44	54.3	490	7	ADH59248	Adh59248	Human	sec	319	44	54.3	490	8	ADG56843	Adg56843	Novel	hum
247	44	54.3	490	7	ADH55240	Adh55240	Novel	hum	320	44	54.3	490	8	ADG55739	Adg55739	Novel	hum
248	44	54.3	490	7	ADH55792	Adh55792	Novel	hum	321	44	54.3	490	8	ADG58499	Adg58499	Novel	hum
249	44	54.3	490	7	ADI38027	Adi38027	Human	sec	322	44	54.3	490	8	ADG70865	Adg70865	Novel	hum
250	44	54.3	490	7	ADI64011	Adi64011	Novel	hum	323	44	54.3	490	8	ADG92655	Adg92655	Human	sec
251	44	54.3	490	7	ADI63459	Adi63459	Novel	hum	324	44	54.3	490	8	ADG57947	Adg57947	Novel	hum
252	44	54.3	490	7	ADH81873	Adh81873	Novel	hum	325	44	54.3	490	8	ADG53531	Adg53531	Novel	hum
253	44	54.3	490	7	ADH81321	Adh81321	Novel	hum	326	44	54.3	490	8	ADG71417	Adg71417	Novel	hum
254	44	54.3	490	7	ADJ26295	Adj26295	Human	sec	327	44	54.3	490	8	ADG81604	Adg81604	Human	PRO
255	44	54.3	490	7	ADM04936	Adm04936	Human	pro	328	44	54.3	490	8	ADH30566	Adh30566	Human	PRO
256	44	54.3	490	7	ADM82490	Adm82490	Novel	hum	329	44	54.3	490	8	ADH11933	Adh11933	Novel	hum
257	44	54.3	490	7	ADN15889	Adn15889	Novel	hum	330	44	54.3	490	8	ADG52355	Adg52355	Novel	hum
258	44	54.3	490	7	ADN16518	Adn16518	Novel	hum	331	44	54.3	490	8	ADG81052	Adg81052	Human	PRO
259	44	54.3	490	7	ADN15337	Adn15337	Novel	hum	332	44	54.3	490	8	ADG56291	Adg56291	Novel	hum
260	44	54.3	490	7	ADN14785	Adn14785	Novel	hum	333	44	54.3	490	8	ADH12557	Adh12557	Novel	hum
261	44	54.3	490	7	ADN14960	Adn14960	Novel	hum	334	44	54.3	490	8	ADG61403	Adg61403	Novel	hum
262	44	54.3	490	8	ADC81047	Adc81047	Novel	hum	335	44	54.3	490	8	ADH28490	Adh28490	Human	PRO
263	44	54.3	490	8	ADE79210	Ade79210	Human	sec	336	44	54.3	490	8	ADG54635	Adg54635	Novel	hum
264	44	54.3	490	8	ADD76495	Add76495	Human	PRO	337	44	54.3	490	8	ADG59675	Adg59675	Novel	hum
265	44	54.3	490	8	ADD87859	Add87859	Human	PRO	338	44	54.3	490	8	ADH20444	Adh20444	Human	sec
266	44	54.3	490	8	ADD86263	Add86263	Human	PRO	339	44	54.3	490	8	ADH07299	Adh07299	Human	sec
267	44	54.3	490	8	ADE79634	Ade79634	Human	sec	340	44	54.3	490	8	ADH59844	Adh59844	Human	sec
268	44	54.3	490	8	ADE75711	Ade75711	Human	PRO	341	44	54.3	490	8	ADH06872	Adh06872	Human	sec
269	44	54.3	490	8	ADE73310	Ade73310	Human	sec	342	44	54.3	490	8	ADI81099	Adi81099	Human	PRO
270	44	54.3	490	8	ADE23287	Ade23287	Human	PRO	343	44	54.3	490	8	ADI18614	Adi18614	Human	sec
271	44	54.3	490	8	ADE23839	Ade23839	Human	PRO	344	44	54.3	490	8	ADI65334	Adi65334	Human	sec
272	44	54.3	490	8	ADE24482	Ade24482	Human	PRO	345	44	54.3	490	8	ADI37597	Adi37597	Human	sec
273	44	54.3	490	8	ADD87307	Add87307	Human	PRO	346	44	54.3	490	8	ADG09842	Adg09842	Novel	hum
274	44	54.3	490	8	ADE89173	Ade89173	Human	PRO	347	44	54.3	490	8	ADH97393	Adh97393	Human	sec
275	44	54.3	490	8	ADE73845	Ade73845	Human	sec	348	44	54.3	490	8	ADI15313	Adi15313	Novel	hum
276	44	54.3	490	8	ADE18312	Ade18312	Human	PRO	349	44	54.3	490	8	ADG09190	Adg09190	Human	sec
277	44	54.3	490	8	ADE88621	Ade88621	Human	PRO	350	44	54.3	490	8	ADI65761	Adi65761	Human	sec
278	44	54.3	490	8	ADE99399	Ade99399	Human	sec	351	44	54.3	490	8	ADI14645	Adi14645	Novel	hum
279	44	54.3	490	8	ADE94641	Ade94641	Human	PRO	352	44	54.3	490	8	ADH60504	Adh60504	Human	sec
280	44	54.3	490	8	ADE91052	Ade91052	Human	PRO	353	44	54.3	490	8	ADI18240	Adi18240	Novel	hum
281	44	54.3	490	8	ADE95193	Ade95193	Human	PRO	354	44	54.3	490	8	ADJ99561	Adj99561	Human	sec
282	44	54.3	490	8	ADE93303	Ade93303	Human	PRO	355	44	54.3	490	8	ADL08754	Adl08754	Human	sec
283	44	54.3	490	8	ADF34884	Adf34884	Human	PRO	356	44	54.3	490	8	ADM25095	Adm25095	Human	sec
284	44	54.3	490	8	ADE98518	Ade98518	Human	sec	357	44	54.3	490	8	ADJ63521	Adj63521	Novel	hum
285	44	54.3	490	8	ADE92199	Ade92199	Novel	hum	358	44	54.3	490	8	ADM29845	Adm29845	Human	sec
286	44	54.3	490	8	ADE90500	Ade90500	Human	PRO	359	44	54.3	490	8	Adj77416	Adj77416	Human	PRO
287	44	54.3	490	8	ADE91647	Ade91647	Novel	hum	360	44	54.3	490	8	Adj65538	Adj65538	Human	PRO
288	44	54.3	490	8	ADE98945	Ade98945	Human	sec	361	44	54.3	490	8	Adm27674	Adm27674	Human	PRO
289	44	54.3	490	8	ADG40415	Adg40415	Human	sec	362	44	54.3	490	8	Adm42398	Adm42398	Human	PRO
290	44	54.3	490	8	ADF73809	Adf73809	Human	sec	363	44	54.3	490	8	Ado06167	Ado06167	Human	PRO
291	44	54.3	490	8	ADG02226	Adg02226	Human	PRO	364	44	54.3	490	8	ADM28260	Adm28260	Human	PRO
292	44	54.3	490	8	ADG22012	Adg22012	Novel	hum	365	44	54.3	490	8	ADR11019	Adr11019	Human	sec
293	44	54.3	490	8	ADG20082	Adg20082	Human	PRO	366	44	54.3	490	8	ADI95742	Adi95742	Human	PRO
294	44	54.3	490	8	ADF97988	Adf97988	Human	PRO	367	44	54.3	490	8	ADI96294	Adi96294	Novel	hum
295	44	54.3	490	8	ADG24205	Adg24205	Novel	hum	368	44	54.3	490	8	ADS74567	Ads74567	Novel	hum
296	44	54.3	490	8	ADF98559	Adf98559	Human	PRO	369	44	54.3	490	8	ADS32246	Ads32246	Novel	hum
297	44	54.3	490	8	ADG03390	Adg03390	Human	PRO	370	44	54.3	490	8	ADT03230	Adt03230	Human	PRO
298	44	54.3	490	8	ADF99111	Adf99111	Human	PRO	371	44	54.3	490	8	ADT03604	Adt03604	Human	sec
299	44	54.3	490	8	ADG16696	Adg16696	Human	PRO	372	44	54.3	490	8	ADZ03281	Adz03281	Human	sec
300	44	54.3	490	8	ADG05155	Adg05155	Human	PRO	373	44	54.3	490	9	AEA37842	Aea37842	Human	sec
301	44	54.3	490	8	ADG19422	Adg19422	Human	PRO	374	44	54.3	490	9	AEBl4027	Aeb14027	Cancer	ce
302	44	54.3	490	8	ADF73385	Adf73385	Human	sec	375	44	54.3	490	7	ADE28643	Ade28643	Human	NOV
303	44	54.3	490	8	ADG13259	Adg13259	Human	PRO	376	44	54.3	549	8	ADM93388	Adm93388	Human	NOV
304	44	54.3	490	8	ADG08316	Adg08316	Novel	hum	377	44	54.3	549	8	ADU01093	Adu01093	Human	pro
305	44	54.3	490	8	ADG15486	Adg15486	Human	PRO	378	44	54.3	671	8	ADU15475	Adul5475	Novel	hum
306	44	54.3	490	8	ADF96884	Adf96884	Human	PRO	379	44	54.3	671	8	ADJ32011	Adj32011	Plant	ino
307	44	54.3	490	8	ADG06069	Adg06069	Human	PRO	380	43	53.1	164	8	ADJ32013	Aay59420	Catalpa	i
308	44	54.3	490	8	ADG23653	Adg23653	Novel	hum	381	43	53.1	232	3	AAJ59420	Aay59420	Plant	ino
309	44	54.3	490	8	ADG03942	Adg03942	Human	PRO	382	43	53.1	270	8	ADJ32011	Aay59420	Catalpa	i
310	44	54.3	490	8	ADG24843	Adg24843	Novel	hum	383	43	53.1	347	7	ADC13496	Adc13496	Protein	w
311	44	54.3	490	8	ADG07140	Adg07140	Novel	hum	384	43	53.1	351	3	AAJ59418	Aay59418	Corn	inos
312	44	54.3	490	8	ADG07692	Adg07692	Novel	hum	385	43	53.1	351	7	ADC13498	Adc13498	Protein	w
313	44	54.3	490	8	ADG55187	Adg55187	Novel	hum	386	43	53.1	569	2	AAJ37113	Aay37113	Chlamydia	
314	44	54.3	490	8	ADG60851	Adg60851	Novel	hum	387	42	51.9	18	2	AAW09593	Aaw09593	Thrombopo	
315	44	54.3	490	8	ADG61955	Adg61955	Novel	hum	388	42	51.9	18	2	AAW36744	Aaw36744	Thrombopo	
316	44	54.3	490	8	ADG92228	Adg92228	Human	sec	389	42	51.9	18	4	AAU25963	Aau25963	Human	thr

390	42	51.9	145	8	ADJ32024	Adj32024 Plant ino	463	39	48.1	416	3	AAB53130	Aab53130 Macaca mu
391	42	51.9	159	3	AAy59431	Aay59431 Wheat ino	464	39	48.1	465	8	ADx77013	Adx77013 Plant ful
392	42	51.9	176	8	ADJ32015	Adj32015 Plant ino	465	39	48.1	484	9	ADx80686	Adx80686 Channel c
393	42	51.9	202	8	ADJ32026	Adj32026 Plant ino	466	39	48.1	533	6	ABU24808	Abu24808 Protein e
394	42	51.9	226	8	ADJ32019	Adj32019 Plant ino	467	39	48.1	591	4	ABG29309	Abg29309 Novel hum
395	42	51.9	244	3	AAy59422	Aay59422 Rice inos	468	39	48.1	765	6	ABU44799	Abu44799 Protein e
396	42	51.9	281	3	AAy59433	Aay59433 Wheat ino	469	39	48.1	791	8	ABM84773	Abm84773 Human dia
397	42	51.9	338	3	AAy59426	Aay59426 Soybean i	470	39	48.1	816	4	ABG25783	Abg25783 Novel hum
398	42	51.9	338	5	ABb92858	Abb92858 Herbicida	471	39	48.1	816	4	ABG04035	Abg04035 Novel hum
399	42	51.9	348	7	ADc13494	Adc13494 Protein w	472	39	48.1	831	4	AAM39264	Aam39264 Human pol
400	42	51.9	353	8	ADT56492	Adt56492 Plant pol	473	39	48.1	831	8	ADQ20117	Adq20117 Human sof
401	42	51.9	391	8	ADT55670	Adt55670 Plant pol	474	39	48.1	831	8	ADQ91460	Adq91460 Amino aci
402	42	51.9	403	8	ADx76324	Adx76324 Plant ful	475	39	48.1	831	9	ADY70286	Ady70286 Human bet
403	42	51.9	439	8	ADN20330	Adn20330 Bacterial	476	39	48.1	831	9	ADY70628	Ady70628 Human pge
404	42	51.9	789	7	ADm25991	Adm25991 Hyperther	477	39	48.1	833	8	ADO71748	Ado71748 BFLP1698
405	41.5	51.2	113	7	ADc00869	Adc00869 Enterohae	478	39	48.1	838	4	AAM41050	Aam41050 Human pol
406	41	50.6	14	2	AAW09542	Aaw09542 Thrombopo	479	39	48.1	840	8	ABM84774	Abm84774 Human dia
407	41	50.6	14	2	AAW36693	Aaw36693 Thrombopo	480	39	48.1	845	4	ABG21852	Abg21852 Novel hum
408	41	50.6	14	4	AAU25912	Aau25912 Human thr	481	39	48.1	880	8	ADL46155	Adl46155 Murine so
409	41	50.6	59	7	ABM73985	Abm73985 DNA clone	482	39	48.1	891	6	ABU49944	Abu49944 Protein e
410	41	50.6	82	5	ABP33075	Abp33075 Human ORF	483	39	48.1	892	4	AAU38250	Aau38250 Salmonell
411	41	50.6	100	8	ADs10566	Ads10566 Human the	484	39	48.1	892	6	ABU47096	Abu47096 Protein e
412	41	50.6	198	7	ADb80946	Adb80946 RING-SH c	485	39	48.1	892	6	ABU48057	Abu48057 Protein e
413	41	50.6	198	9	ADx07064	Adx07064 Cyclin-de	486	39	48.1	898	7	ABO63805	Abo63805 Klebsiell
414	41	50.6	240	8	ABM81851	Abm81851 Tumour-as	487	39	48.1	908	8	ADO71747	Ado71747 BFLP1698
415	41	50.6	470	2	AAW25766	Aaw25766 Human CAR	488	39	48.1	1019	8	ABM81255	Abm81255 Tumour-as
416	41	50.6	470	3	AAy98167	Aay98167 Human TRA	489	39	48.1	1019	8	ADN18867	Adn18867 Bacterial
417	41	50.6	470	6	ABO07168	Abo07168 Human p53	490	39	48.1	1049	8	ADN18867	Adn18867 Bacterial
418	41	50.6	470	7	ADb80943	Adb80943 RING-SH c	491	39	48.1	1162	8	ADO71742	Ado71742 BFLP1698
419	41	50.6	470	8	ADP54388	Adp54388 Human PRO	492	39	48.1	1162	8	ADO71735	Ado71735 BFLP1698
420	41	50.6	470	9	ADx07062	Adx07062 Cyclin-de	493	39	48.1	1162	8	ADO71738	Ado71738 BFLP1698
421	41	50.6	470	9	ADY19800	Ady19800 PRO polyp	494	39	48.1	1162	8	ADO71744	Ado71744 BFLP1698
422	41	50.6	1885	8	ADL98343	Adl98343 SSL-1, SE	495	39	48.1	1162	8	ADO71737	Ado71737 BFLP1698
423	40	49.4	75	3	AAG27112	Aag27112 Zea may	496	39	48.1	1162	8	ADO71740	Ado71740 BFLP1698
424	40	49.4	113	8	ADM87627	Adm87627 Human EST	497	39	48.1	1162	8	ADO71739	Ado71739 BFLP1698
425	40	49.4	143	3	AAG16560	Aag16560 Arabidops	498	39	48.1	1162	8	ADO71741	Ado71741 BFLP1698
426	40	49.4	185	9	AEb35749	Aeb35749 L. pneumo	499	39	48.1	1162	8	ADO71743	Ado71743 BFLP1698
427	40	49.4	297	4	ABb12473	Abb12473 Human bon	500	39	48.1	1162	8	ADO71733	Ado71733 Human BFL
428	40	49.4	365	7	ADf58880	Adf58880 Human pol	501	39	48.1	1162	8	ADO71736	Ado71736 BFLP1698
429	40	49.4	374	7	ADc13502	Adc13502 Protein w	502	39	48.1	1162	8	ADO71745	Ado71745 BFLP1698
430	40	49.4	375	6	ABU31595	Abu31595 Protein e	503	39	48.1	1162	8	ADO71734	Ado71734 BFLP1698
431	40	49.4	384	8	ADY24810	Ady24810 Plant ful	504	39	48.1	1588	6	ABR52925	Abr52925 Protein s
432	40	49.4	391	7	ABO64064	Abo64064 Klebsiell	505	39	48.1	1588	7	ADK62502	Adk62502 Disease t
433	40	49.4	406	2	AAW70326	Aaw70326 Secreted	506	39	48.1	1588	8	ADN18982	Adn18982 Bacterial
434	40	49.4	427	9	ABM96829	Abm96829 M. xanthu	507	38.5	47.5	1858	5	ABB90859	Abb90859 Herbicida
435	40	49.4	461	6	ABR53263	Abr53263 Protein s	508	38	46.9	14	2	AAW09538	Aaw09538 Thrombopo
436	40	49.4	461	7	ADK63638	Adk63638 Disease t	509	38	46.9	14	2	AAW36689	Aaw36689 Thrombopo
437	40	49.4	461	8	ADs43875	Ads43875 Bacterial	510	38	46.9	14	4	AAU25908	Aau25908 Human thr
438	40	49.4	468	8	ADM87165	Adm87165 Human pro	511	38	46.9	116	7	ABO75047	Abo75047 Pseudomon
439	40	49.4	497	6	ABP71435	Abp71435 B. halodu	512	38	46.9	143	3	AAB25206	Aab25206 Eucalyptu
440	40	49.4	497	8	ADs28233	Ads28233 Bacterial	513	38	46.9	225	4	AAG98805	Aag98805 Human cel
441	40	49.4	521	8	ADI30117	Adi30117 M. methyl	514	38	46.9	225	9	AEb36789	Aeb36789 L. pneumo
442	40	49.4	569	5	ABb94300	Abb94300 Chlamydia	515	38	46.9	225	9	AEb40164	Aeb40164 L. pneumo
443	40	49.4	772	2	AAr511478	Aar511478 Human TLE	516	38	46.9	240	4	AAM00973	Aam00973 Human bon
444	40	49.4	772	2	AAr51111	Aar51111 Human TLE	517	38	46.9	254	3	AAy83791	Aay83791 S. fradia
445	40	49.4	772	8	ADn04243	Adn04243 Antipsori	518	38	46.9	262	6	ABJ26540	Abj26540 Aspergill
446	40	49.4	1044	6	ADA13361	Ada13361 Human int	519	38	46.9	262	8	ADR85983	Adr85983 Aspergill
447	40	49.4	1093	8	ADO44174	Ado44174 Structura	520	38	46.9	277	1	AAp70298	Aap70298 Sequence
448	40	49.4	1101	4	AAW50136	Aam50136 Human GTP	521	38	46.9	277	2	AAW33275	Aaw33275 S. fradia
449	40	49.4	1101	6	ADA13349	Ada13349 Human int	522	38	46.9	289	4	AAG98803	Aag98803 Human cel
450	39.5	48.8	113	7	ADc00042	Adc00042 Enterohae	523	38	46.9	311	4	AAG98800	Aag98800 Human cel
451	39.5	48.8	113	7	ADc00340	Adc00340 Enterohae	524	38	46.9	330	4	AAU48008	Aau48008 Propionib
452	39.5	48.8	113	7	ADc00493	Adc00493 Enterohae	525	38	46.9	330	6	ABM44527	Abm44527 Propionib
453	39	48.1	14	2	AAW09536	Aaw09536 Thrombopo	526	38	46.9	338	4	AAG98797	Aag98797 Human cel
454	39	48.1	14	2	AAW09544	Aaw09544 Thrombopo	527	38	46.9	375	4	AAG98796	Aag98796 Human cel
455	39	48.1	14	2	AAW36687	Aaw36687 Thrombopo	528	38	46.9	411	6	ABP79669	Abp79669 N. gonorr
456	39	48.1	14	2	AAW36695	Aaw36695 Thrombopo	529	38	46.9	419	4	AAG98793	Aag98793 Human cel
457	39	48.1	14	2	AAW36768	Aaw36768 Thrombopo	530	38	46.9	425	4	AAU35805	Aau35805 Helicobac
458	39	48.1	14	4	AAU25906	Aau25906 Human thr	531	38	46.9	425	6	ABU30955	Abu30955 Protein e
459	39	48.1	14	4	AAU25914	Aau25914 Human thr	532	38	46.9	447	6	ABU47669	Abu47669 Protein e
460	39	48.1	184	8	ADx67893	Adx67893 Plant ful	533	38	46.9	447	6	ABU44845	Abu44845 Protein e
461	39	48.1	218	2	AAW20717	Aaw20717 H. pylori	534	38	46.9	464	7	ADB70147	Abd70147 C. neofor
462	39	48.1	334	8	ADY12323	Ady12323 Plant ful	535	38	46.9	469	6	ABU40299	Abu40299 Protein e

536	38	46.9	469	6	ABU41961	Abu41961 Protein e
537	38	46.9	471	6	ABU15562	Abu15562 Protein e
538	38	46.9	471	8	ADS15046	AdS15046 Pseudomon
539	38	46.9	482	4	AAU27892	Aau27892 Human con
540	38	46.9	488	4	AAG98787	Aag98787 Human cel
541	38	46.9	490	4	AAG98786	Aag98786 Human cel
542	38	46.9	522	8	ADT57198	Adt57198 Plant pol
543	38	46.9	530	8	ADY07757	Ady07757 Plant ful
544	38	46.9	555	7	ABO79567	Abo79567 Pseudomon
545	38	46.9	578	4	AAU27720	Aau27720 Human ful
546	38	46.9	604	6	ADA48262	Ada48262 Rice prot
547	38	46.9	604	7	ABM86602	Abm86602 Rice abio
548	38	46.9	625	4	AAAB95823	Aab95823 Human pro
549	38	46.9	764	2	AAW94359	Aaw94359 Human fac
550	38	46.9	893	6	ABU22898	Abu22898 Protein e
551	38	46.9	1243	8	ADR09787	Adr09787 Human pro
552	38	46.9	3262	3	AAAY53675	Aay53675 Mechanica
553	38	46.9	3264	3	AAAY53676	Aay53676 Protein 2
554	38	46.9	5183	6	AAE14793	Aae14793 Human mic
555	38	46.9	5183	8	ADO44006	Ado44006 Amino aci
556	38	46.9	5183	9	ADX06154	Adx06154 Cyclin-de
557	37.5	46.3	12	2	AAW09524	Aaw09524 Thrombopo
558	37.5	46.3	12	2	AAW36675	Aaw36675 Thrombopo
559	37.5	46.3	12	3	AAAB16993	Aab16993 TPO-mimet
560	37.5	46.3	12	4	AAU25894	Aau25894 Human thr
561	37.5	46.3	12	5	ABB72879	Abb72879 TPO mimet
562	37.5	46.3	12	7	ADJ73030	Adj73030 TPO mimet
563	37.5	46.3	12	8	ADJ52665	Adj52665 CH1 delet
564	37.5	46.3	12	8	ADJ51626	Adj51626 CH1 delet
565	37.5	46.3	16	2	AAW09464	Aaw09464 Thrombopo
566	37.5	46.3	16	2	AAW33329	Aaw33329 Thrombopo
567	37.5	46.3	16	3	AAAB17019	Aab17019 TPO-mimet
568	37.5	46.3	16	4	AAU25829	Aau25829 Human thr
569	37.5	46.3	16	5	ABB72905	Abb72905 TPO mimet
570	37.5	46.3	16	7	ADJ73057	Adj73057 TPO mimet
571	37.5	46.3	16	8	ADJ52692	Adj52692 CH1 delet
572	37.5	46.3	16	8	ADJ51653	Adj51653 CH1 delet
573	37.5	46.3	221	7	ABO70412	Abo70412 Pseudomon
574	37.5	46.3	262	6	ABU15700	Abu15700 Protein e
575	37.5	46.3	282	5	AAU72889	Aau72889 Human met
576	37.5	46.3	351	8	ADS98200	Ads98200 Protein f
577	37	45.7	12	2	AAW09581	Aaw09581 Thrombopo
578	37	45.7	12	2	AAW09582	Aaw09582 Thrombopo
579	37	45.7	12	2	AAW36733	Aaw36733 Thrombopo
580	37	45.7	12	4	AAU25952	Aau25952 Human thr
581	37	45.7	12	4	AAU25951	Aau25951 Human thr
582	37	45.7	13	2	AAW36732	Aaw36732 Thrombopo
583	37	45.7	57	4	AAAM83901	Aam83901 Human imm
584	37	45.7	107	6	ABP75804	Abp75804 Human sec
585	37	45.7	126	3	AAG02882	Aag02882 Human sec
586	37	45.7	141	8	ADX74883	Adx74883 Plant ful
587	37	45.7	156	4	ABG19580	Abg19580 Novel hum
588	37	45.7	162	4	AAAM90349	Aam90349 Human imm
589	37	45.7	173	6	ABU23115	Abu23115 Protein e
590	37	45.7	218	8	ADQ65547	Adq65547 Novel hum
591	37	45.7	220	7	ABO80735	Abo80735 Pseudomon
592	37	45.7	223	4	ABB11685	Abb11685 Human sec
593	37	45.7	248	6	ABU40125	Abu40125 Protein e
594	37	45.7	273	6	ABM68436	Abm68436 Photorhab
595	37	45.7	277	8	ADX78585	Adx78585 Plant ful
596	37	45.7	291	5	ABB75700	Abb75700 Human het
597	37	45.7	291	7	ADM04528	Adm04528 Human pro
598	37	45.7	304	5	ABB97526	Abb97526 Novel hum
599	37	45.7	308	6	ABU40110	Abu40110 Protein e
600	37	45.7	312	4	ABG13680	Abg13680 Novel hum
601	37	45.7	327	4	ABG16297	Abg16297 Novel hum
602	37	45.7	333	3	AAAY97409	Aay97409 Zebrafish
603	37	45.7	354	7	ABO78447	Abo78447 Pseudomon
604	37	45.7	366	8	ADJ48863	Adj48863 Oil-assoc
605	37	45.7	384	8	ADY11295	Ady11295 Plant ful
606	37	45.7	394	8	ADX91505	Adx91505 Plant ful
607	37	45.7	412	6	ADA36849	Ada36849 Acinetoba
608	37	45.7	417	2	AAW42401	Aaw42401 TRAF2 bin

609	37	45.7	455	6	ABU23370	Abu23370
610	37	45.7	471	6	ABU21814	Abu21814
611	37	45.7	482	5	ADK34554	Adk34554 Novel hum
612	37	45.7	491	6	ABU22810	Abu22810 Protein e
613	37	45.7	530	4	AAG90434	Aag90434 C glutami
614	37	45.7	541	5	ABB57264	Abb57264 Mouse isc
615	37	45.7	542	7	ADD48911	Add48911 Rat Prote
616	37	45.7	570	6	ABU62313	Abu62313 Rat MINT
617	37	45.7	741	4	ABG19581	Abg19581 Novel hum
618	37	45.7	749	4	ABG24612	Abg24612 Novel hum
619	37	45.7	750	6	ABU62312	Abu62312 Rat MINT
620	37	45.7	750	6	ABU62311	Abu62311 Rat MxX2
621	37	45.7	766	4	ABG13668	Abg13668 Novel hum
622	37	45.7	773	5	AAE28503	Aae28503 Maize suc
623	37	45.7	1026	5	ABP73754	Abp73754 Candida a
624	37	45.7	1027	2	AAW17057	Aaw17057 Candida a
625	37	45.7	1072	7	ABM85376	Abm85376 Mouse pro
626	37	45.7	1115	5	AAAB71230	Aab71230 Human leg
627	37	45.7	1115	7	ABW01535	Abw01535 Human lgs
628	37	45.7	1115	8	ADJ71905	Adj71905 Human lgs
629	37	45.7	1244	8	ADN22906	Adn22906 Bacterial
630	37	45.7	1244	8	ADN22907	Adn22907 Bacterial
631	37	45.7	1244	8	ADN22908	Adn22908 Bacterial
632	37	45.7	1566	4	ABB64864	Abb64864 Drosophil
633	37	45.7	1577	8	ABM83607	Abm83607 Human dia
634	37	45.7	1617	8	ABM83606	Abm83606 Human dia
635	37	45.7	1653	6	ABR41799	Abr41799 Human DIT
636	37	45.7	1653	6	ABU11522	Abu11522 Human MDD
637	37	45.7	1682	4	ABG13667	Abg13667 Novel hum
638	37	45.7	1682	4	ABG07239	Abg07239 Novel hum
639	37	45.7	1911	2	AAAR99534	Aar99534 Dermatomy
640	37	45.7	1912	4	AAAM39288	Aam39288 Human pol
641	37	45.7	1912	6	AAE36110	Aae36110 Human chr
642	37	45.7	1912	6	AAE36109	Aae36109 Human chr
643	37	45.7	1912	8	ADP12552	Adp12552 Protein e
644	37	45.7	1912	8	ABM81343	Abm81343 Tumour-as
645	37	45.7	1936	4	AAAM41074	Aam41074 Human pol
646	37	45.7	2159	4	ABG13664	Abg13664 Novel hum
647	37	45.7	2386	2	AAW13153	Aaw13153 S. pombe
648	37	45.7	2386	8	ADN19726	Adn19726 Bacterial
649	37	45.7	2485	2	AAAR59922	Aar59922 RAS assoc
650	37	45.7	2485	2	AAAR59921	Aar59921 RAS assoc
651	37	45.7	2503	4	ABG13554	Abg13554 Novel hum
652	37	45.7	2503	4	ABG07238	Abg07238 Novel hum
653	37	45.7	2531	5	ADI16935	Adi16935 Rat NOX
654	37	45.7	2531	5	ADI16934	Adi16934 Rat NOX
655	37	45.7	2531	7	ADEG3713	Ade63713 Rat Prote
656	37	45.7	2531	7	ADEG3705	Ade63705 Rat Prote
657	37	45.7	2531	7	ADEG3709	Ade63709 Rat Prote
658	37	45.7	2531	7	ADEG3701	Ade63701 Rat Prote
659	37	45.7	2818	2	AAAR22268	Aar22268 Nf1 gene
660	37	45.7	2818	2	AAW13280	Aaw13280 Human neu
661	37	45.7	2818	2	AAE05486	Aae05486 Human neu
662	37	45.7	2818	5	ABB08078	Abb08078 Human neu
663	37	45.7	2871	4	ABG13665	Abg13665 Novel hum
664	37	45.7	2871	4	ABG13555	Abg13555 Novel hum
665	37	45.7	2871	4	ABG07243	Abg07243 Novel hum
666	37	45.7	2891	4	ABG13666	Abg13666 Novel hum
667	37	45.7	2891	4	ABG07240	Abg07240 Novel hum
668	37	45.7	2891	4	ABG13556	Abg13556 Novel hum
669	36.5	45.1	10	2	AAW09557	Aaw09557 Thrombopo
670	36.5	45.1	10	2	AAW36708	Aaw36708 Thrombopo
671	36.5	45.1	10	4	AAU25927	Aau25927 Human thr
672	36.5	45.1	316	9	AEB27292	Aeb27292 Pinus rad
673	36.5	45.1	348	4	AAAB84996	Aab84996 Human nov
674	36.5	45.1	348	5	ABP51981	Abp51981 Human tra
675	36.5	45.1	348	8	ADO32262	Ado32262 Novel hum
676	36.5	45.1	350	2	AAW56265	Aaw56265 Mouse rec
677	36.5	45.1	350	3	AAAY91933	Aay91933 Murine Re
678	36.5	45.1	379	5	ABB92530	Abb92530 Herbicida
679	36.5	45.1	860	5	ABB07689	Abb07689 Rat gluta
680	36	44.4	14	2	AAW09539	Aaw09539 Thrombopo
681	36	44.4	14	2	AAW36691	Aaw36691 Thrombopo

Abu23370	Protein e
Abu21814	Protein e
Adk34554	Novel hum
Abu22810	Protein e
Aag90434	C glutami
Abb57264	Mouse isc
Add48911	Rat Prote
Abu62313	Rat MINT
Abg19581	Novel hum
Abg24612	Novel hum
Abu62312	Rat MINT
Abu62311	Rat MxX2
Abg13668	Novel hum
Aae28503	Maize suc
Abp73754	Candida a
Aaw17057	Candida a
Abm85376	Mouse pro
Aab71230	Human leg
Abw01535	Human lgs
Adj71905	Human lgs
Adn22906	Bacterial
Adn22907	Bacterial
Adn22908	Bacterial
Abb64864	Drosophil
Abm83607	Human dia
Abm83606	Human dia
Abr41799	Human DIT
Abu11522	Human MDD
Abg13667	Novel hum
Abg07239	Novel hum
Aar99534	Dermatomy
Aam39288	Human pol
Aae36110	Human chr
Aae36109	Human chr
Adp12552	Protein e
Abm81343	Tumour-as
Aam41074	Human pol
Abg13664	Novel hum
Aaw13153	S. pombe
Adn19726	Bacterial
Aar59922	RAS assoc
Aar59921	RAS assoc
Abg13554	Novel hum
Abg07238	Novel hum
Adi16935	Rat NOX
Adi16934	Rat NOX
Ade63713	Rat Prote
Ade63705	Rat Prote
Ade63709	Rat Prote
Ade63701	Rat Prote
Aar22268	Nf1 gene
Aaw13280	Human neu
Aae05486	Human neu
Abb08078	Human neu
Abg13665	Novel hum
Abg13555	Novel hum
Abg07243	Novel hum
Abg13666	Novel hum
Abg07240	Novel hum
Abg13556	Novel hum
Aaw09557	Thrombopo
Aaw36708	Thrombopo
Aau25927	Human thr
Aeb27292	Pinus rad
Aab84996	Human nov
Abp51981	Human tra
Ado32262	Novel hum
Aaw56265	Mouse rec
Aay91933	Murine Re
Abb92530	Herbicida
Abb07689	Rat gluta
Aaw09539	Thrombopo
Aaw36691	Thrombopo

682	36	44.4	14	2	AAW36690	Aaw36690	Thrombopo
683	36	44.4	14	2	AAW36769	Aaw36769	Thrombopo
684	36	44.4	14	3	AAB16996	Aab16996	TPO-mimet
685	36	44.4	14	4	AAU25909	Aau25909	Human thr
686	36	44.4	14	4	AAU25910	Aau25910	Human thr
687	36	44.4	14	5	ABB72882	Abb72882	TPO mimet
688	36	44.4	14	7	ADJ73033	Adj73033	TPO mimet
689	36	44.4	14	8	ADJ52668	Adj52668	CH1 delet
690	36	44.4	14	8	ADJ51629	Adj51629	CH1 delet
691	36	44.4	85	4	AAU56464	Aau56464	Propionib
692	36	44.4	85	6	ABM52983	Abm52983	Propionib
693	36	44.4	99	3	AAG01004	Aag01004	Human sec
694	36	44.4	140	7	ABO69232	Abo69232	Pseudomon
695	36	44.4	157	4	AAM15531	Aam15531	Peptide #
696	36	44.4	157	4	ABB34538	Abb34538	Peptide #
697	36	44.4	157	4	AAM28021	Aam28021	Peptide #
698	36	44.4	157	4	ABB29364	Abb29364	Peptide #
699	36	44.4	157	4	ABB19947	Abb19947	Protein #
700	36	44.4	157	4	AAM67721	Aam67721	Human bon
701	36	44.4	157	4	AAM55324	Aam55324	Human bra
702	36	44.4	157	4	ABG49359	Abg49359	Human liv
703	36	44.4	157	4	AAM03280	Aam03280	Peptide #
704	36	44.4	159	2	AAR21745	Aar21745	LECGNA8.
705	36	44.4	159	2	AAR83114	Aar83114	Snowdrop
706	36	44.4	162	4	AAG73460	Aag73460	Human gen
707	36	44.4	162	5	ABB90156	Abb90156	Human pol
708	36	44.4	183	4	AAG72613	Aag72613	Human OR-
709	36	44.4	187	7	ABO73960	Abo73960	Pseudomon
710	36	44.4	198	8	ADX79170	Adx79170	Plant ful
711	36	44.4	203	6	ABP99502	Abp99502	Human sec
712	36	44.4	203	6	ABR00975	Abro00975	Human gen
713	36	44.4	203	6	ADA98177	Ada98177	Human sec
714	36	44.4	203	6	ADA44031	Ada44031	Human sec
715	36	44.4	203	7	ADC20347	Adc20347	Human sec
716	36	44.4	203	7	ADF10687	Adf10687	Human sec
717	36	44.4	204	3	AAB51840	Aab51840	Human sec
718	36	44.4	205	3	AAB06132	Aab06132	Caenorhab
719	36	44.4	221	4	AAB66068	Aab66068	Human TAN
720	36	44.4	221	6	ABO32722	Abo32722	Secreted
721	36	44.4	221	8	ADQ10398	Adq10398	Human pol
722	36	44.4	221	8	ADQ98270	Adq98270	Human pol
723	36	44.4	230	7	ADB69992	Adb69992	C. neofor
724	36	44.4	239	8	ADY22640	Ady22640	Plant ful
725	36	44.4	250	8	ADG22312	Adg22312	Cyanophag
726	36	44.4	255	7	ADC33028	Adc33028	Human nov
727	36	44.4	275	5	ABP56265	Abp56265	Human NIP
728	36	44.4	275	8	ADW39231	Adw39231	Human BNI
729	36	44.4	282	2	AAR11871	Aar11871	Recombina
730	36	44.4	282	8	ABM81515	Abm81515	Tumour-as
731	36	44.4	282	9	ADM08788	Adm08788	Human pro
732	36	44.4	282	9	AEA17049	Aea17049	Alzheimer
733	36	44.4	287	2	AAW51096	Aaw51096	Ehrlichia
734	36	44.4	287	2	AAU06960	Aau06960	E. canis
735	36	44.4	287	3	AAU71481	Aay71481	Ehrlichia
736	36	44.4	287	3	AAB36190	Aab36190	Ehrlichia
737	36	44.4	287	4	AAU04200	Aau04200	Variable
738	36	44.4	287	5	AAU96104	Aau96104	Ehrlichia
739	36	44.4	287	5	ABG77951	Abg77951	Ehrlichia
740	36	44.4	287	6	ADA09767	Ada09767	E. canis
741	36	44.4	287	9	ADM04260	Adm04260	Ehrlichia
742	36	44.4	295	7	ADM27021	Adm27021	Hyperther
743	36	44.4	298	8	ADY09054	Ady09054	Plant ful
744	36	44.4	302	6	ABP71305	Abp71305	Murine on
745	36	44.4	303	6	ABJ37060	Abj37060	Human bre
746	36	44.4	303	6	ABP71304	Abp71304	Human onc
747	36	44.4	303	8	ADQ18331	Adq18331	Human sof
748	36	44.4	317	7	ADR14487	Adrl4487	Human NF-
749	36	44.4	317	7	ADC97305	Adc97305	E. faeciu
750	36	44.4	320	3	AAG10428	Aagl0428	Arabidops
751	36	44.4	322	7	ADF04707	Adf04707	Bacterial
752	36	44.4	335	2	AAR99247	Aar99247	Mouse int
753	36	44.4	335	2	AAW44005	Aaw44005	Murine in
754	36	44.4	335	2	AAW56710	Aaw56710	p35 subun

36	44.4	335	4	AAAB36650	Aab36650	Mouse cyt
36	44.4	335	4	AAE11196	Aae11196	Mouse int
36	44.4	335	4	AAE11194	Aae11194	Mouse int
36	44.4	335	4	AAE11197	Aae11197	Mouse int
36	44.4	335	6	ABU10116	Abu10116	Mouse int
36	44.4	335	7	ADC07182	Adc07182	Mouse cyt
36	44.4	335	8	ADS00012	Ads00012	Murine in
36	44.4	335	9	ADW97245	Adw97245	Murine p4
36	44.4	335	9	ADZ65017	Adz65017	Mouse p40
36	44.4	344	5	AAU77548	Aau77548	Human mem
36	44.4	346	7	ADC31361	Adc31361	Human nov
36	44.4	350	8	ADT60598	Adt60598	Plant pol
36	44.4	351	2	AAAY35573	Aay35573	C. pneumo
36	44.4	351	6	ABU27028	Abu27028	Protein e
36	44.4	353	8	ADX93161	Adx93161	Plant ful
36	44.4	354	4	AAAB72416	Aab72416	Murine DN
36	44.4	354	4	AAE13014	Aae13014	Mouse deo
36	44.4	354	4	ABG27268	Abg27268	Novel hum
36	44.4	357	7	ADB64961	Adb64961	Human pro
36	44.4	357	8	ADW39229	Adw39229	Human BNI
36	44.4	364	3	AAG48975	Aag48975	Arabidops
36	44.4	364	5	ABB93221	Abb93221	Herbicida
36	44.4	369	8	ADQ65073	Adq65073	Novel hum
36	44.4	371	3	AAG48974	Aag48974	Arabidops
36	44.4	390	4	AAAB66067	Aab66067	Human TAN
36	44.4	390	6	ABO32721	Abo32721	Secreted
36	44.4	390	8	ADQ10397	Adq10397	Human pol
36	44.4	390	8	ADQ98269	Adq98269	Human pol
36	44.4	397	3	AAAY74889	Aay74889	Neisseria
36	44.4	397	6	ABU37997	Abu37997	Protein e
36	44.4	405	7	ADE08159	Ade08159	Novel pro
36	44.4	405	7	ADH87580	Adh87580	Enterococ
36	44.4	407	3	AAG48973	Aag48973	Arabidops
36	44.4	409	4	AAE11931	Aae11931	Human CG1
36	44.4	409	8	ADS10855	Ads10855	Human the
36	44.4	423	4	AAAB66065	Aab66065	Human TAN
36	44.4	423	6	ABU08369	Abu08369	Amino aci
36	44.4	423	6	ABO32719	Abo32719	Secreted
36	44.4	423	8	ADQ10395	Adq10395	Human pol
36	44.4	423	8	ADQ98267	Adq98267	Human pol
36	44.4	458	8	ADS10856	Ads10856	Human the
36	44.4	465	7	ADE09164	Ade09164	Novel pro
36	44.4	465	8	ADS12105	Ads12105	Human the
36	44.4	465	8	ADS12106	Ads12106	Human the
36	44.4	466	4	ABH22527	Abh22527	Human tra
36	44.4	479	9	ABB70033	Abb70033	Drosophil
36	44.4	480	4	ABB66251	Abb66251	Eucalyptu
36	44.4	480	4	ABB70230	Abb70230	Drosophil
36	44.4	497	5	AAAM49570	Aam49570	B. napus
36	44.4	497	7	ADC25951	Adc25951	Oilseed r
36	44.4	513	8	ADN20604	Adn20604	Bacterial
36	44.4	515	6	ADB10524	Adb10524	Alloiococ
36	44.4	531	8	ADQ14471	Adq14471	Mouse hyp
36	44.4	531	8	ADS73606	Ads73606	Mouse IL-
36	44.4	538	7	ABO73681	Abo73681	Pseudomon
36	44.4	541	6	ABU89759	Abu89759	Protein d
36	44.4	541	7	ADD48913	Add48913	Human pro
36	44.4	541	8	ADR09709	Adr09709	Human pro
36	44.4	555	6	ABB99660	Abb99660	Amino aci
36	44.4	569	5	ABP40727	Abp40727	Staphyloc
36	44.4	569	8	ADS08025	Ads08025	Staphyloc
36	44.4	572	8	ADG31207	Adg31207	Novel mou
36	44.4	572	8	ADQ67668	Adq67668	Novel hum
36	44.4	573	4	AAAB97498	Aab97498	Murine DN
36	44.4	595	6	ABU22647	Abu22647	Protein e
36	44.4	595	6	ABU19799	Abu19799	Protein e
36	44.4	595	6	ABU21867	Abu21867	Protein e
36	44.4	597	3	AAG30765	Aag30765	Arabidops
36	44.4	617	7	ADC31032	Adc31032	Human nov
36	44.4	619	8	ADQ67755	Adq67755	Novel hum
36	44.4	621	4	AAAB94049	Aab94049	Human pro
36	44.4	638	7	ADK40958	Adk40958	Novel hum

828	36	44.4	638	8	ADR15677	Adr15677 Kinase 63	901	35	43.2	12	2	AAW09574	Aaw09574 Thrombopo
829	36	44.4	642	6	ADB10522	Adb10522 Alloiococ	902	35	43.2	12	4	AAW36725	Aaw36725 Thrombopo
830	36	44.4	645	3	AAG45573	Aag45573 Arabidops	903	35	43.2	12	4	Aau25944	Aau25944 Human thr
831	36	44.4	648	7	AAE38322	Aae38322 Human mem	904	35	43.2	14	2	AAW09543	Aaw09543 Thrombopo
832	36	44.4	651	8	ADJ49850	Adj49850 Oil-assoc	905	35	43.2	14	2	AAW36765	Aaw36765 Thrombopo
833	36	44.4	667	5	ADR05219	Adr05219 Human tes	906	35	43.2	14	2	AAW36694	Aaw36694 Thrombopo
834	36	44.4	693	7	AAE38320	Aae38320 Human mem	907	35	43.2	14	3	AAB16998	Aab16998 TPO-mimet
835	36	44.4	700	6	ABP56624	Abp56624 Human MTS	908	35	43.2	14	4	AAU25913	Aau25913 Human thr
836	36	44.4	700	7	ADI10414	Adi10414 Human cel	909	35	43.2	14	5	ABB72884	Abb72884 TPO mimet
837	36	44.4	700	8	ADJ46938	Adj46938 Human tra	910	35	43.2	14	7	ADJ73035	Adj73035 TPO mimet
838	36	44.4	701	8	ADP04659	Adp04659 Sea squir	911	35	43.2	14	8	ADJ52670	Adj52670 CH1 delet
839	36	44.4	706	7	AAE38321	Aae38321 Human mem	912	35	43.2	14	8	ADJ51631	Adj51631 CH1 delet
840	36	44.4	713	7	ADE07937	Ade07937 Novel pro	913	35	43.2	27	5	AAU70544	Aau70544 Carcinoem
841	36	44.4	762	3	AAG30764	Aag30764 Arabidops	914	35	43.2	27	7	ADG44703	Adg44703 CEA bindi
842	36	44.4	764	3	AAG30763	Aag30763 Arabidops	915	35	43.2	29	7	ABW00978	Abw00978 Mutant JA
843	36	44.4	764	5	ABB92076	Abb92076 Herbicida	916	35	43.2	29	7	Ade86421	Ade86421 Mutant JA
844	36	44.4	796	3	AAB06067	Aab06067 Caenorhab	917	35	43.2	30	4	AAO08982	Aao08982 Human pol
845	36	44.4	797	2	ADF77527	Adf77527 Caenorhab	918	35	43.2	35	6	ABU61407	Abu61407 Low densi
846	36	44.4	810	3	AAG45572	Aag45572 Arabidops	919	35	43.2	35	6	ABU61267	Abu61267 Human A d
847	36	44.4	812	3	AAG45571	Aag45571 Arabidops	920	35	43.2	35	5	ADP21489	Adp21489 Human LDL
848	36	44.4	821	7	ADB74284	Adb74284 Mycobacte	921	35	43.2	51	5	AAE15311	Aae15311 GmSUTr1 cy
849	36	44.4	843	5	AAU77552	Aau77552 Hman memb	922	35	43.2	52	4	AAU57514	Aau57514 Propionib
850	36	44.4	843	7	AAE38319	Aae38319 Human mem	923	35	43.2	52	6	ABM54033	Abm54033 Propionib
851	36	44.4	850	5	AAU82750	Aau82750 Amino aci	924	35	43.2	58	4	AAW92725	Aam92725 Human dig
852	36	44.4	850	5	AAU74757	Aau74757 Human pro	925	35	43.2	74	2	AAW38398	Aay38398 Human sec
853	36	44.4	858	2	ADF77528	Adf77528 Caenorhab	926	35	43.2	78	4	AAW85706	Aam85706 Human imm
854	36	44.4	858	3	AAB06068	Aab06068 Caenorhab	927	35	43.2	86	3	AAW94967	Aay94967 Human sec
855	36	44.4	878	6	ABM69030	Abm69030 Photorhab	928	35	43.2	103	5	ABB48769	Abb48769 Listeria
856	36	44.4	887	6	ABU40971	Abu40971 Protein e	929	35	43.2	104	7	ADB64569	Abb64569 Human pro
857	36	44.4	891	4	AAU34554	Aau34554 E. coli c	930	35	43.2	124	3	AAW53612	Aab53612 Human col
858	36	44.4	891	6	ABU28609	Abu28609 Protein e	931	35	43.2	124	3	AAG00384	Aag00384 Human sec
859	36	44.4	892	2	ADF77529	Adf77529 Caenorhab	932	35	43.2	136	4	AAU63009	Aau63009 Propionib
860	36	44.4	892	3	AAB06069	Aab06069 Caenorhab	933	35	43.2	140	4	ABM59528	Abm59528 Propionib
861	36	44.4	894	7	ADF04439	Adf04439 Bacterial	934	35	43.2	140	6	ABM59528	Abm59528 Propionib
862	36	44.4	928	2	AAR77399	Aar77399 BHV1 gi g	935	35	43.2	148	4	AAW94560	Aam94560 Human rep
863	36	44.4	933	2	AAR27807	Aar27807 Bovine he	936	35	43.2	148	5	ABG61739	Abg61739 Novel ova
864	36	44.4	933	2	AAR41343	Aar41343 Bovine he	937	35	43.2	149	4	ABB61925	Abb61925 Drosophil
865	36	44.4	938	5	ABB47445	Abb47445 Listeria	938	35	43.2	179	7	ADF04394	Adf04394 Bacterial
866	36	44.4	1059	8	ADN22521	Adn22521 Bacterial	939	35	43.2	179	7	ADH86006	Adh86006 Enterococ
867	36	44.4	1059	8	ADN22522	Adn22522 Bacterial	940	35	43.2	180	6	ADA33903	Ada33903 Acinetoba
868	36	44.4	1318	8	ADQ66591	Adq66591 Novel hum	941	35	43.2	191	6	ABM69595	Abm69595 Photorhab
869	35.5	43.8	54	1	AAP91239	Aap91239 Epidermal	942	35	43.2	204	7	ABO68038	Abo68038 Pseudomon
870	35.5	43.8	74	4	ABG11153	Abg11153 Novel hum	943	35	43.2	209	2	AAW43259	Aay43259 H. pylori
871	35.5	43.8	74	4	ABG00620	Abg00620 Novel hum	944	35	43.2	210	2	AAW98657	Aaw98657 H. pylori
872	35.5	43.8	230	6	ABU41706	Abu41706 Protein e	945	35	43.2	210	3	AAB14302	Aab14302 Human sec
873	35.5	43.8	253	3	AAB42375	Aab42375 Human ORF	946	35	43.2	210	4	AAW85224	Aab85224 Human sec
874	35.5	43.8	311	4	ABB63010	Abb63010 Drosophil	947	35	43.2	214	4	AAO13786	Aao13786 Human pol
875	35.5	43.8	339	4	AAG70840	Aag70840 C albican	948	35	43.2	214	4	AEA20855	Aea20855 Novel hum
876	35.5	43.8	353	8	ADT60388	Adt60388 Plant pol	949	35	43.2	247	6	ABU18579	Abu18579 Protein e
877	35.5	43.8	360	8	ADY22374	Ady22374 Plant ful	950	35	43.2	256	6	ABU36441	Abu36441 Protein e
878	35.5	43.8	488	7	ADM04834	Adm04834 Human pro	951	35	43.2	256	6	ABB84616	Abb84616 M. tuberc
879	35.5	43.8	546	4	ABB65311	Abb65311 Drosophil	952	35	43.2	259	9	AEA19898	Aea19898 Novel hum
880	35.5	43.8	755	4	AAB94435	Aab94435 Human pro	953	35	43.2	293	9	ABE35765	Aeb35765 L. pneumo
881	35.5	43.8	755	5	ABP69413	Abp69413 Human pol	954	35	43.2	293	9	ABE39215	Aeb39215 L. pneumo
882	35.5	43.8	823	2	AAW25119	Aay25119 Human MDC	955	35	43.2	295	4	ABB52831	Abb52831 Escherich
883	35.5	43.8	823	9	ADX06478	Adx06478 Cyclin-de	956	35	43.2	296	5	ABB98136	Abb98136 Human PMM
884	35.5	43.8	838	4	ABB61967	Abb61967 Drosophil	957	35	43.2	303	4	ABG08486	Abg08486 Novel hum
885	35.5	43.8	859	2	AAW25118	Aay25118 Human MDC	958	35	43.2	303	6	ADA33592	Ada33592 Acinetoba
886	35.5	43.8	859	9	ADX06474	Adx06474 Amino aci	959	35	43.2	308	8	ADF66220	Adf66220 Rat Ian5(
887	35.5	43.8	867	2	AAW30208	Aay30208 Human alp	960	35	43.2	312	8	ADS28334	Ads28334 Bacterial
888	35.5	43.8	870	6	ABR63662	Abx06476 Cyclin-de	961	35	43.2	313	4	AAW73003	Aag73003 Olfactory
889	35.5	43.8	870	9	ADX06476	Adx06476 Cyclin-de	962	35	43.2	314	7	ADM11499	Adm11499 Murine ol
890	35.5	43.8	893	8	ADQ66143	Adq66143 Novel hum	963	35	43.2	314	7	ADM25896	Adm25896 Hyperther
891	35.5	43.8	899	7	ADE71110	Ade71110 ADAM22, S	964	35	43.2	315	2	AAW48792	Aaw48792 Homo sapi
892	35.5	43.8	916	8	ADM74169	Adm74169 Human NOV	965	35	43.2	315	6	ABO07185	Ab007185 Human p53
893	35.5	43.8	1088	5	ABJ05495	Abj05495 Human bre	966	35	43.2	315	7	ADE61338	Ade61338 Human Pro
894	35.5	43.8	1088	5	ABJ01044	Abj01044 Human bre	967	35	43.2	315	7	ADG14302	Adg14302 Human Spr
895	35.5	43.8	1097	5	ABP73762	Abp73762 Candida a	968	35	43.2	315	8	ADG47834	Adg47834 Human spr
896	35.5	43.8	1412	8	ADJ50150	Adj50150 Oil-assoc	969	35	43.2	315	8	ADG47837	Adg47837 Mouse spr
897	35.5	43.8	1566	6	ABR59716	Abx59716 Human RER	970	35	43.2	315	8	ADJ95401	Adj95401 Human spr
898	35.5	43.8	1666	6	ABJ25895	Abj25895 Aspergill	971	35	43.2	315	8	ADJ95404	Adj95404 Mouse spr
899	35.5	43.8	1680	6	ABJ26495	Abj26495 Aspergill	972	35	43.2	315	8	ADN11324	Adn11324 Human spr
900	35	43.2	11	5	ABP47661	Abp47661 N. mening	973	35	43.2	315	8	ADQ21111	Adq21111 Human sof

974 35 43.2 315 8 ADP23634 Adp23634 PRO polyp
975 35 43.2 315 8 ADS15841 Ads15841 Human spr
976 35 43.2 315 9 ADY19896 Ady19896 PRO polyp
977 35 43.2 315 9 ADY15230 Ady15230 PRO polyp
978 35 43.2 319 4 AAB76547 Aab76547 Corynebac
979 35 43.2 321 4 AAB76546 Aab76546 Corynebac
980 35 43.2 321 4 AAG91586 Aag91586 C glutami
981 35 43.2 321 7 ADL65571 Adl65571 C. glutam
982 35 43.2 321 8 ADS41652 Ads41652 Bacterial
983 35 43.2 322 2 AAR47859 Aar47859 Human LDL
984 35 43.2 322 2 AAY41120 Aay41120 Soybean s
985 35 43.2 322 2 AAY22599 Aay22599 LDL recep
986 35 43.2 329 6 ABU23480 Abu23480 Protein e
987 35 43.2 335 2 AAW80661 Aaw80661 S. pneumo
988 35 43.2 339 4 ABB58379 Abb58379 Drosophil
989 35 43.2 339 8 ADX88353 Adx88353 Plant ful
990 35 43.2 343 4 AAG72916 Aag72916 Human olf
991 35 43.2 346 4 AAU58400 Aau58400 Propionib
992 35 43.2 346 6 ABM54919 Abm54919 Propionib
993 35 43.2 351 8 ADM41707 Adm41707 Protein m
994 35 43.2 356 2 AAR48547 Aar48547 Sequence
995 35 43.2 357 6 ABM68597 Abm68597 Photorhab
996 35 43.2 363 7 ADJ68958 Adj68958 Human hea
997 35 43.2 367 8 ADP04625 Adp04625 Sea squir
998 35 43.2 377 6 ABU41138 Abu41138 Protein e
999 35 43.2 385 7 ADF07153 Adf07153 Bacterial
1000 35 43.2 394 8 ADX67615 Adx67615 Plant ful

ALIGNMENTS

RESULT 1
AAW09535
ID AAW09535 standard; protein; 14 AA.
XX
AC AAW09535;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 28; 106pp; English.
XX

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 81; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CSLQEFLSHGGYVC 14
Db 1 CSLQEFLSHGGYVC 14
RESULT 2
AAW36686
ID AAW36686 standard; peptide; 14 AA.
XX
AC AAW36686;
XX
DT 27-FEB-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 28; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 81; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
Db 1 CSLQEFLSHGGYVC 14

RESULT 3
AAB16997
ID AAB16997 standard; peptide; 14 AA.
XX
AC AAB16997;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:53.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 213; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
Db 1 CSLQEFLSHGGYVC 14

RESULT 4
AAU25905
ID AAU25905 standard; peptide; 14 AA.
XX
AC AAU25905;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #91.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 21; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk.The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
Db 1 CSLQEFLSHGGYVC 14

Db	1	CSLQEFLSHGGYVC 14	
RESULT 7			
ADJ52669			
ID	ADJ52669	standard; peptide; 14 AA.	
XX			
AC	ADJ52669;		
XX			
DT	06-MAY-2004	(first entry)	
XX			
DE	CH1	deleted mimetibody-related peptide SeqID488.	
XX			
KW	CH1	deleted mimetibody; immunosuppressive; cardiovascular; cardiant; hypotensive; neuroprotective; nootropic; antibacterial; virucide; fungicide; gene therapy; immune disorder; cardiovascular disease; arrhythmia; hypertension; heart failure; neurodegenerative; multiple sclerosis; dementia; Alzheimer's disease; anaemia; cancerous condition; infectious disease; bacterial infection; viral infection; fungal infection.	
KW			
KW			
KW			
KW			
KW			
KW			
XX			
OS	Unidentified.		
OS	Synthetic.		
XX			
PN	WO2004002417-A2.		
XX			
PD	08-JAN-2004.		
XX			
PF	27-JUN-2003; 2003WO-US020347.		
XX			
PR	28-JUN-2002; 2002US-0392431P.		
XX			
PA	(CENZ) CENTOCOR INC.		
XX			
PI	Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC; Kutoloski KA;		
PI			
DR	WPI; 2004-082870/08.		
XX			
XX	New CH1-deleted mimetibody polypeptides and nucleic acids, useful for modulating, treating, alleviating, preventing an immune, cardiovascular, or neurodegenerative disease or disorder, anemia, cancer, or infectious diseases.		
PS	Claim 2; SEQ ID NO 488; 129pp; English.		
XX			
CC	This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an immunosuppressive, cardiovascular, cardiant, hypotensive, neuroprotective, nootropic, antibacterial, virucide or fungicide activity. In addition, the disclosed sequences may prove useful for gene therapy. The CH1-deleted mimetibody is useful for diagnosing or treating a disease condition in a cell, tissue, organ or animal, specifically for modulating, treating, alleviating, preventing the incidence or reducing the symptoms of an immune, cardiovascular (for example arrhythmia, hypertension or heart failure), or neurodegenerative (for example multiple sclerosis, dementia or Alzheimer's disease) diseases or disorders, anaemia, cancerous conditions, or infectious diseases (for example bacterial, viral or fungal infection). The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.		
XX			
SQ	Sequence 14 AA;		
Query Match 100.0%; Score 81; DB 8; Length 14;			
Best Local Similarity 100.0%; Pred. No. 2.8e-06;			
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CSLQEFLSHGGYVC 14	
Db	1	CSLQEFLSHGGYVC 14	

RESULT 8			
ADJ51630			
ID	ADJ51630	standard; peptide; 14 AA.	
XX			
AC	ADJ51630;		
XX			
DT	06-MAY-2004	(first entry)	
XX			
DE	CH1	deleted mimetibody-related peptide SeqID488.	
XX			
KW	CH1	deleted mimetibody; osteopathic; cardiovascular-Gen; dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen; gynaecological-Gen; hepatotropic; haemostatic; immunomodulator; antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic; ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor; TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder; dental disorder; oral disorder; dermatological disorder; ear disorder; nose disorder; throat disorder; endocrine disorder; metabolic disorder; gastrointestinal disorder; gynaecological disorder; hepatic disorder; obstetric disorder; haematologic disorder; immunological disorder; allergic disorder; infectious disorder; musculoskeletal disorder; oncological disorder; neurological disorder; nutritional disorder; ophthalmologic disorder; pediatric disorder; psychiatric disorder; renal disorder; pulmonary disorder.	
XX			
OS	Unidentified.		
OS	Synthetic.		
XX			
PN	WO2004002424-A2.		
XX			
PD	08-JAN-2004.		
XX			
PF	30-JUN-2003; 2003WO-US020495.		
XX			
PR	28-JUN-2002; 2002US-0392431P.		
PR	19-SEP-2002; 2002US-0412144P.		
XX			
PA	(CENZ) CENTOCOR INC.		
XX			
PI	Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC; Kutoloski KA;		
XX			
DR	WPI; 2004-082872/08.		
XX			
PT	New CH1 deleted mimetibody polypeptide and nucleic acid, useful for diagnosing, preventing or treating cardiovascular, dermatologic, endocrine, gastrointestinal, gynecologic, infectious, neurologic and nutritional disorders.		
PT			
PT			
PT			
XX			
PS	Claim 14; SEQ ID NO 488; 123pp; English.		
XX			
CC	This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an osteopathic, cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen, gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic, immunomodulator, antiallergic, muscular-Gen, cytostatic, antiinflammatory, neuroleptic, ophthalmological, nephrotropic or respiratory-Gen activity acting as a tumour necrosis factor (TNF)-modulator or cytokine-agonist. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of diseases or conditions associated with aberrant expression or activity of the CH1 deleted mimetibody, such as a bone or joint, cardiovascular, dental or oral, dermatological, ear, nose or throat, endocrine, metabolic, gastrointestinal, gynaecological, hepatic, obstetric, haematologic, immunological, allergic, infectious, musculoskeletal, oncological, neurological, nutritional, ophthalmologic, pediatric, psychiatric, renal or pulmonary disorders. The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.		
XX			
SQ	Sequence 14 AA;		

XX										
SQ	Sequence 344 AA;									
Query Match		59.3%;		Score 48;		DB 3;		Length 344;		
Best Local Similarity		61.5%;		Pred. No. 21;						
Matches		8;		Conservative		2;		Mismatches		
						3;		Indels		
								0;		
								Gaps		
								0;		
QY	2 SLQEFLSHGGYVC 14									
	:									
Db	139 SLEDFLYHEGYAC 151									
RESULT 11										
AAW81503										
ID	AAW81503 standard; protein; 660 AA.									
XX										
AC	AAW81503;									
XX										
DT	02-FEB-1999 (first entry)									
XX										
DE	Dead Box Y (DBY) gene product.									
XX										
KW	Non-recombining region; human; Y chromosome; X homologue; testis; DBY;									
KW	infertility; sperm; gene alteration; inhibitor; Dead Box Y.									
XX										
OS	Homo sapiens.									
XX										
PN	WO9846747-A2.									
XX										
PD	22-OCT-1998.									
XX										
PF	10-APR-1998; 98WO-US007115.									
XX										
PR	11-APR-1997; 97US-0041877P.									
XX										
PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.									
XX										
PI	Lahn BT, Page DC;									
XX										
DR	WPI; 1998-568729/48.									
DR	N-PSDB; AAV69633.									
XX										
PT	Novel genes in the non-combining region of Y chromosome - useful to									
PT	diagnose if male infertility or reduced sperm count has a genetic basis.									
XX										
PS	Claim 6; Fig 3A-B; 54pp; English.									
XX										
CC	This represents the amino acid sequence of the Dead Box Y (DBY) gene									
CC	product. The invention relates to genes occurring on the non-recombining									
CC	region of the human Y chromosome. The sequences fall into two classes:									
CC	(1) X-homologous DNA which are expressed in many organs, having									
CC	functional X homologues and (2) testis-specific DNA sequences. Y									
CC	chromosomal DNA from males with known conditions such as infertility and									
CC	reduced sperm count can be assessed using the invention to determine									
CC	whether the condition is associated with or caused by the occurrence of									
CC	the gene or gene alteration. Candidate inhibitors of the enzymatic									
CC	activity of the genes can be assessed using in vitro assays									
XX										
SQ	Sequence 660 AA;									
Query Match		59.3%;		Score 48;		DB 2;		Length 660;		
Best Local Similarity		61.5%;		Pred. No. 41;						
Matches		8;		Conservative		2;		Mismatches		
						3;		Indels		
								0;		
								Gaps		
								0;		
QY	2 SLQEFLSHGGYVC 14									
	:									
Db	454 SLEDFLYHEGYAC 466									
RESULT 12										
ABO07237										
ID	ABO07237 standard; protein; 660 AA.									
XX										

AC	ABO07237;									
XX										
DT	13-AUG-2003 (first entry)									
XX										
DE	Human p53 modifying protein, SEQ ID 197.									
XX										
KW	Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;									
KW	antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;									
KW	lung cancer; ovarian cancer; angiogenesis; cell cycle;									
KW	apoptotic disorder; cell proliferation disorder.									
XX										
OS	Homo sapiens.									
XX										
PN	WO200299122-A1.									
XX										
PD	12-DEC-2002.									
XX										
PF	03-JUN-2002; 2002WO-US017382.									
XX										
PR	05-JUN-2001; 2001US-0296076P.									
PR	10-OCT-2001; 2001US-0328605P.									
PR	15-FEB-2002; 2002US-0357253P.									
XX										
PA	(EXEL-) EXELIXIS INC.									
XX										
PI	Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;									
XX										
DR	WPI; 2003-156859/15.									
DR	N-PSDB; ACD13410.									
XX										
PT	Identifying modulators of the p53 pathway for use in treating apoptotic									
PT	or cell proliferation disorders, comprises screening for agents that									
PT	modulate activity of a human ortholog of genes that modify the p53									
PT	pathway in Drosophila.									
XX										
PS	Example 2; Page 567-569; 678pp; English.									
XX										
CC	The invention relates to identifying (M1) a candidate p53 pathway									
CC	modulating agent, by contacting an assay system comprising a purified HM									
CC	polypeptide (human orthologue of genes that modify the p53 pathway in									
CC	Drosophila) or nucleic acid with a test agent under conditions, where but									
CC	for the presence of the test agent, the system provides a reference									
CC	activity, and detecting a test agent-biased activity of the assay system.									
CC	Also included are modulating (M2) a p53 pathway of a cell (comprising									
CC	contacting a cell defective in p53 function with a candidate modulator									
CC	that specifically binds to a HM polypeptide comprising an HM amino acid									
CC	sequence, where p53 function is restored), modulating (M3) a p53 pathway									
CC	in a mammalian cell (comprising contacting the cell with an agent that									
CC	specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)									
CC	a disease in a patient (comprising: (a) obtaining a biological sample									
CC	from the patient; (b) contacting the sample with a probe for HM									
CC	expression; (c) comparing the results with a control; and (d) determining									
CC	whether the comparison indicates a likelihood disease). (M1) is useful									
CC	for identifying modulators of the p53 pathway. A probe for HM expression									
CC	is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,									
CC	in a patient, where the cancer has greater than 25 % expression level.									
CC	Modulators identified by (M1) are useful in a variety of diagnostic and									
CC	therapeutic applications, where disease or disorder prognosis is related									
CC	to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell									
CC	proliferation disorders (e.g. cancer). Another two new methods (M2 and									
CC	M3) are useful for modulating the p53 pathway of a cell, thus restoring									
CC	the p53 function of the cell, so that the cell undergoes normal									
CC	proliferation or progression through the cell cycle. (M2) and (M3) are									
CC	also useful for treating defects in the p53 pathway such as angiogenic,									
CC	apoptotic or cell proliferation disorders. The present sequence									
CC	represents a human p53 pathway modifying protein									
..										
XX	Sequence 660 AA;									
SO										

QY 2 SLQEFLSHGGYVC 14
Db 454 SLEDFLYHEGYAC 466

RESULT 13
ADL13058
ID ADL13058 standard; protein; 660 AA.
XX
AC ADL13058;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell protein #121.
XX
KW Hepatototropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
DR WPI; 2004-068610/07.
XX
PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Disclosure; SEQ ID NO 787; 141pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 660 AA;

Query Match 59.3%; Score 48; DB 8; Length 660;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
Db 454 SLEDFLYHEGYAC 466

RESULT 14

ADM33393
ID ADM33393 standard; protein; 660 AA.
XX
AC ADM33393;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human PRO84194 protein SEQ ID NO:10.
XX
KW human; PRO; antipsoriatic; gene therapy; psoriasis.
XX
OS Homo sapiens.
XX
PN WO2004028447-A2.
XX
PD 08-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US027382.
XX
PR 26-SEP-2002; 2002US-0414484P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305092/28.
DR N-PSDB; ADM33392.
XX
PT New PRO nucleic acid or polypeptide e.g., PRO37523, PRO71267, PRO71295 or
PT PRO1843, useful for preparing a composition for diagnosing or treating
PT psoriasis in a mammal.
XX
PS Claim 9; SEQ ID NO 10; 143pp; English.
XX
CC The present invention describes a novel human isolated nucleic acid (I)
CC designated PRON, where n represents a 4 or 5 digit number. Also
CC described: (1) a vector comprising (1); (2) a host cell comprising the
CC vector; (3) a process for producing a PRO polypeptide; (4) an isolated
CC PRO polypeptide; (5) a chimeric molecule comprising the polypeptide fused
CC to a heterologous amino acid sequence; (6) an antibody which specifically
CC binds to a PRO polypeptide; (7) a composition of matter comprising the
CC polypeptide or its agonist or antagonist or antibody and a carrier; (8)
CC an article of manufacture comprising a container and its label and the
CC composition of matter, contained within the container, where label on the
CC container indicates that the composition of matter can be used for
CC treating an immune related disorder; (9) a method of treating psoriasis
CC in a mammal; (10) a method for determining the presence of a PRO
CC polypeptide in a sample suspected of containing the polypeptide; (11) a
CC method of identifying a compound that inhibits the expression of a gene
CC encoding a PRO polypeptide from the present invention; (12) a method of
CC identifying a compound that mimics the activity of a PRO polypeptide from
CC the present invention; and (13) a method of diagnosing psoriasis in a
CC mammal. The PRO sequences have antipsoriatic activity, and can be used in
CC gene therapy. The PRO nucleic acid or polypeptide sequences from the
CC present invention can be used for preparing a composition for diagnosing
CC or treating psoriasis. The present sequence represents a human PRO
CC protein from the present invention.
XX
SQ Sequence 660 AA;

Query Match 59.3%; Score 48; DB 8; Length 660;
Best Local Similarity 61.5%; Pred. NO. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
Db 454 SLEDFLYHEGYAC 466

RESULT 15
ADW08765
ID ADW08765 standard; protein; 660 AA.

XX AC ADW08765;
XX XX
DT 24-MAR-2005 (first entry)
XX DE Human protein which is up-regulated in HCV-infected tissue - SEQ ID 102.
XX KW gene targeting; hepatitis C virus infection; protein deactivation;
KW protein activation.
XX OS Homo sapiens.
XX XX
FN EP1493750-A2.
XX XX
PD 05-JAN-2005.
XX XX
PF 28-JUN-2004; 2004EP-00015098.
XX PR 30-JUN-2003; 2003GB-00015248.
XX XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX PI Berndt P, Kilby PM, Rugman P;
XX XX
DR WPI; 2005-050476/06.
XX XX
PT New targets for an antiviral compound having at least one down- and up-
PT regulated hepatitis C virus (HCV) polypeptide, useful in predicting
PT outcomes, treating or preventing HCV infections.
XX XX
PS Claim 1; SEQ ID NO 102; 346pp; English.
XX CC The invention comprises the amino acid sequences of protein targets for
CC anti-hepatitis C virus (anti-HCV) drugs. The human protein targets of the
CC invention are either up-regulated or down-regulated in HCV-infected
CC tissue. The proteins of the invention are useful in the preparation of a
CC medicament for the treatment or prevention of HCV infection. The present
CC amino acid sequence represents a human protein of the invention which is
CC up-regulated in HCV-infected tissue.
XX XX
SQ Sequence 660 AA;

Query Match 59.3%; Score 48; DB 9; Length 660;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYYVC 14
||:|||||
Db 454 SLEDFLYHEGYAC 466

RESULT 16
AEA24054
ID AEA24054 standard; protein; 660 AA.
XX AC
XX AEA24054;
XX XX
DT 11-AUG-2005 (first entry)
XX XX
DE Human PRO polypeptide SEQ ID NO 596.
XX XX
KW immune disorder; PRO; Antiinflammatory; Dermatological;
KW Immunosuppressive; Antirheumatic; Antiarthritic; Osteopathic;
KW Muscular-Gen.; Vasotropic; Antianemic; Antipsoriatic; Immunostimulant.
XX OS Homo sapiens.
XX XX
PN WO2005051988-A2.
XX XX
PD 09-JUN-2005.
XX XX
PF 02-MAR-2004; 2004WO-US006460.
XX XX

PR 03-MAR-2003; 2003US-0451884P.
XX (GETH) GENENTECH INC.
XX XX
PI Abbas A, Bodary S, Clark H, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX XX
DR WPI; 2005-417958/42.
DR N-PSDB; AEA24053.
XX XX
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, autoimmune hemolytic anemia, or
PT psoriasis.
XX XX
PS Disclosure; SEQ ID NO 596; 966pp; English.
XX XX
CC The invention relates to an isolated nucleic acid. The polypeptide,
CC compound or composition, and methods are useful for diagnosing and
CC treating an immune related disorder, e.g. systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,
CC spondyloarthropathies, systemic sclerosis, idiopathic inflammatory
CC myopathies, Sjogren's syndrome, systemic vasculitis, sarcoidosis,
CC autoimmune hemolytic anemia, autoimmune or immune-mediated skin diseases
CC including bullous skin diseases, erythema multiforme and contact
CC dermatitis, psoriasis, lymphadenopathy, splenomegaly and leukopenia. The
CC present sequence represents the amino acid sequence of a human PRO
CC polypeptide.
XX XX
SQ Sequence 660 AA;

Query Match 59.3%; Score 48; DB 9; Length 660;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYYVC 14
||:|||||
Db 454 SLEDFLYHEGYAC 466

RESULT 17
AAW81501
ID AAW81501 standard; protein; 662 AA.
XX AC
XX AAW81501;
XX XX
DT 02-FEB-1999 (first entry)
XX XX
DE Dead Box X (DBX) gene long transcript amino acid sequence.
XX KW Non-recombining region; human; Y chromosome; X homologue; testis; DBX;
KW infertility; sperm; gene alteration; inhibitor; Dead Box X.
XX OS Homo sapiens.
XX XX
PN WO9846747-A2.
XX PD 22-OCT-1998.
XX PF 10-APR-1998; 98WO-US007115.
XX XX
PR 11-APR-1997; 97US-0041877P.
XX XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX XX
PI Lahn BT, Page DC;
XX XX
DR WPI; 1998-568729/48.
DR N-PSDB; AAV69631.
XX XX
PT Novel genes in the non-combining region of Y chromosome - useful to
PT diagnose if male infertility or reduced sperm count has a genetic basis.
XX XX

PS Disclosure; Fig 3A-B; 54pp; English.

XX This represents the amino acid sequence of the Dead Box X (DBX) gene long

CC transcript. The invention relates to genes occurring on the non-

CC recombining region of the human Y chromosome. The sequences fall into two

CC classes: (1) X-homologous DNA which are expressed in many organs, having

CC functional X homologues and (2) testis-specific DNA sequences. Y

CC chromosomal DNA from males with known conditions such as infertility and

CC reduced sperm count can be assessed using the invention to determine

CC whether the condition is associated with or caused by the occurrence of

CC the gene or gene alteration. Candidate inhibitors of the enzymatic

CC activity of the genes can be assessed using in vitro assays

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 2; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14

Db 456 SLEDFLYHEGYAC 468

RESULT 18

AAW81502

ID AAW81502 standard; protein; 662 AA.

XX

AC AAW81502;

XX

DT 02-FEB-1999 (first entry)

XX

DE Dead Box X (DBX) gene short transcript amino acid sequence.

XX

KW Non-recombining region; human; Y chromosome; X homologue; testis; DBX;

KW infertility; sperm; gene alteration; inhibitor; Dead Box X.

XX

OS Homo sapiens.

XX

PN WO9846747-A2.

XX

PD 22-OCT-1998.

XX

PF 10-APR-1998; 98WO-US007115.

XX

PR 11-APR-1997; 97US-0041877P.

XX

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX

PI Lahn BT, Page DC;

XX

DR WPI; 1998-568729/48.

DR N-PSDB; AAV69632.

XX

PT Novel genes in the non-combining region of Y chromosome - useful to

PT diagnose if male infertility or reduced sperm count has a genetic basis.

XX

PS Disclosure; Fig 3A-B; 54pp; English.

XX

CC This represents the amino acid sequence of the Dead Box X (DBY) gene

CC short transcript. The invention relates to genes occurring on the non-

CC recombining region of the human Y chromosome. The sequences fall into two

CC classes: (1) X-homologous DNA which are expressed in many organs, having

CC functional X homologues and (2) testis-specific DNA sequences. Y

CC chromosomal DNA from males with known conditions such as infertility and

CC reduced sperm count can be assessed using the invention to determine

CC whether the condition is associated with or caused by the occurrence of

CC the gene or gene alteration. Candidate inhibitors of the enzymatic

CC activity of the genes can be assessed using in vitro assays

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 2; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14

Db 456 SLEDFLYHEGYAC 468

RESULT 19

AA96483

ID AA96483 standard; protein; 662 AA.

XX

AC AA96483;

XX

DT 12-SEP-2000 (first entry)

XX

DE Human DBX1, an RNA helicase homologue.

XX

KW Topoisomerase III-alpha; interacting protein; TIII-a IP; RNA helicase;

KW DBX1; nucleosome; supercoiling; chromosome segregation; recombination;

KW stabilization; cell division; apoptosis; cell cycle regulation;

KW cytotoxic; anti-tumour; DBX1.

XX

OS Homo sapiens.

XX

PN WO200032768-A1.

XX

PD 08-JUN-2000.

XX

PF 29-NOV-1999; 99WO-FR002952.

XX

PR 30-NOV-1998; 98FR-00015081.

XX

PA (AVET) AVENTIS PHARMA SA.

XX

PI Fournier A, Goulaouic H, Riou J;

XX

DR WPI; 2000-412316/35.

DR N-PSDB; AAA29207.

XX

PT New nucleic acid encoding ligand for topoisomerase III alpha, useful for

PT inhibiting the enzyme and in drug screening, e.g. for potential

PT anticancer agents.

XX

PS Disclosure; Page 53-56; 68pp; French.

XX

CC This protein is DBX1, which shows homology with RNA helicases but the

CC activity of a helicase has never been demonstrated and its function has

CC not yet been identified. DBX1 possesses the 8 characteristic motifs of

CC the helicases of the family "DEAD". In particular, it appears to be part

CC of the sub-family represented by helicase PL10. The DBX1 gene is situated

CC on the X chromosome and its homologue, which is situated on the Y

CC chromosome has 91 percent identity with a novel protein coding sequence.

CC The novel protein (see AA96482) is a human topoisomerase III-alpha

CC interacting protein (TIII-a IP). TIII-a IP has structural features in

CC common with RNA helicases (e.g. DBX1), which are involved in destruction

CC of the nucleosome structure, supercoiling of DNA, segregation of newly

CC replicated chromosomes, and recombination and stabilization of the

CC genome. Agents that alter interaction between TIII-a IP and TIII-a

CC therefore modulate cell division, replication, transcription,

CC translation, splicing and DNA recombination or repair, so may slow cell

CC growth, block the cell cycle or induce apoptosis. Antibodies and ligands

CC of TIII-a IP are used to prevent, treat or alleviate diseases that

CC involve abnormal regulation of the cell cycle, i.e. they are potential

CC anti-tumour (cytotoxic) agents

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 3; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14

Db 456 SLEDFLYHEGYAC 468

RESULT 20
ABB57045

ID ABB57045 standard; protein; 662 AA.

XX
AC ABB57045;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:68.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99243.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 215-218; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 5; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYVC 14
||:|||||

Db 456 SLEDFLYHEGYAC 468

RESULT 21
ABO07236

ID ABO07236 standard; protein; 662 AA.

XX
AC ABO07236;
XX

DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 196.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-US017382.
PF 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13409.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 565-567; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
XX represents a human p53 pathway modifying protein
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 6; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYVC 14
||:|||||

Db 456 SLEDFLYHEGYAC 468

RESULT 22
ABO52958

ID ABO52958 standard; protein; 662 AA.
XX
AC ABO52958;
XX
DT 09-OCT-2003 (first entry)
XX
DE Human spliceosome associated protein (SAP) #64.
XX
KW Human; SAP; spliceosome associated protein; ribonucleoprotein;
KW RNP complex; RNA affinity substrate; RNP assembly sequence;
KW spliceosomal complex; hnRNP complex; mRNA export complex;
KW mRNA localisation complex; RNA editing complex; intron complex;
KW H complex; telomerase complex; fragile X protein complex;
KW reverse transcriptase complex; gene splicing complex.
XX
OS Homo sapiens.
XX
PN US2003068803-A1.
XX
PD 10-APR-2003.
XX
PF 14-JAN-2002; 2002US-00047991.
XX
PR 12-JAN-2001; 2001US-0261521P.
XX
PA (REED/) REED R.
PA (ZHOU/) ZHOU Z.
XX
PI Reed R, Zhou Z;
XX
DR WPI; 2003-540885/51.
XX
PT Isolating ribonucleoprotein complex, by contacting RNA affinity substrate
PT having ribonucleoprotein assembly sequence and affinity tag, with protein
PT mixture, subjecting complex formed to chromatography, affinity selection.
PS Claim 24; Page; 39pp; English.
XX
CC The invention relates to forming (M1) an isolated ribonucleoprotein (RNP)
CC complex (C), involves contacting an RNA affinity substrate (S) comprising
CC an RNP assembly sequence (AS) and an affinity tag, with a protein mixture
CC to permit formation of (C) on AS, subjecting (C) to chromatographic
CC separation, and subjecting (C) to affinity selection, where the affinity
CC tag (e.g. bacteriophage MS2 coat protein in a fusion protein with E. coli
CC maltose binding protein) binds to an affinity matrix. Also included are
CC an isolated spliceosome preparation (isolated by (M1)), a RNA comprising
CC an RNP complex binding site and at least one phage coat protein
CC recognition site, a nucleic acid encoding the RNA, and treating (M2) a
CC subject having a disorder associated with abnormal RNP complexes (by
CC obtaining a sample of cells from a subject, purifying RNP complexes from
CC the cells of the subject by (M1), determining the presence in the
CC purified RNP complexes of one or more proteins, and normalising the
CC amount of RNPs in the subject. (M1) is useful for forming an isolated RNP
CC complex selected from a spliceosomal complex (selected from E, A, B and C
CC complex), an hnRNP complex, an mRNA export complex, an mRNA localisation
CC complex, an RNA editing complex, an intron complex, or an H complex. (M1)
CC is useful in a diagnostic assay for determining whether a subject has
CC abnormal RNP complexes, (M2) is useful for treating a subject having a
CC disorder associated with abnormal RNP complexes. (M1) is useful for
CC forming an isolated RNP complex such as a telomerase complex, a fragile X
CC protein complex, a reverse transcriptase complex or a gene splicing
CC complex. The present sequence represents a known human spliceosome
CC associated protein (SAP) isolated by the methods of the invention. Note:
CC The present sequence is not shown in the specification but was obtained
CC from Genbank or Swissprot using the information provided in table 1 of
CC the specification
XX
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 6; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||:|||||
Db 456 SLEDFLYHEGYAC 468

RESULT 23
ADN04580

ID ADN04580 standard; protein; 662 AA.
XX
AC ADN04580;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic protein sequence #479.
XX
KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR N-PSDB; ADN04579.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 9; SEQ ID NO 974; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 8; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||:|||||
Db 456 SLEDFLYHEGYAC 468

RESULT 24
ADN03714

ID ADN03714 standard; protein; 662 AA.
XX
AC ADN03714;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic protein sequence #54.
XX

KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX WO2004028479-A2.
PN
XX 08-APR-2004.
PD
XX
XX 25-SEP-2003; 2003WO-US030907.
PF
XX 25-SEP-2002; 2002US-0414006P.
PR
XX (GETH) GENENTECH INC.
PA
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;
PI Wu TD;
PI
XX WPI; 2004-305105/28.
DR N-PSDB; ADN03713.
DR
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
PT
XX
PS Claim 9; SEQ ID NO 108; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 8; Length 662;
Best Local Similarity 61.5%; Pred. NO. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYVC 14
Db ||:|||||
456 SLEDFLYHEGYAC 468

RESULT 25
ADX07534
ID ADX07534 standard; protein; 662 AA.
XX
AC ADX07534;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 2099.
XX
KW cytosstatic; cyclin-dependent kinase; cdk; biomarker.
XX
OS Homo sapiens.
XX
XX WO2005012875-A2.
PN
XX 10-FEB-2005.
PD
XX
XX 29-JUL-2004; 2004WO-US024424.
PF
XX 29-JUL-2003; 2003US-0490890P.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PA
XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
PI
XX WPI; 2005-163068/17.
DR N-PSDB; ADX07533.
DR
XX Biomarkers useful for predicting or determining the response of a mammal
PT

PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 2099; 141pp; English.
XX
CC This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-([5-(1,1-Dimethylethyl)-2-
CC oxazolyl]methylthio)-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence represents a biomarker used in the method of the invention.
XX
SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 9; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYVC 14
Db ||:|||||
456 SLEDFLYHEGYAC 468

RESULT 26
AAG07259
ID AAG07259 standard; protein; 192 AA.
XX
AC AAG07259;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4345.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 45; DB 3; Length 192;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CSLQEFLSHGGYVC 14
|:|: |:|:| |
Db 140 CALKWFVNHGSTVC 153

RESULT 27
ADJ32021
ID ADJ32021 standard; protein; 198 AA.
XX AC ADJ32021;
XX DT 15-APR-2004 (first entry)
XX DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #13.
XX KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.
XX OS Glycine max.
XX PN US2003229917-A1.
XX PD 11-DEC-2003.
XX PF 20-DEC-2002; 2002US-00324316.
XX PR 11-OCT-2000; 2000US-00686521.
XX (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
DR New polynucleotide comprising a sequence encoding a polypeptide having
PT inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
PT physical mapping for plant breeding.
XX
PS Example 3; Fig 1; 32pp; English.
XX
CC This invention relates to a novel isolated polynucleotide which comprises
CC a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
CC kinase activity and comprising a sequence having 80% identity based on
CC the Clustal alignment method with a fully defined sequence comprising 139
CC amino acids, or its complement. The polynucleotide is useful in genetic
CC or physical mapping for plant breeding. The present sequence is that of a

CC plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC exemplification of the invention.
XX
SQ Sequence 198 AA;

Query Match 55.6%; Score 45; DB 8; Length 198;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LQEFLSHGGYV 13
|:|:|:| |
Db 116 LQEFVNHGGVV 126

RESULT 28
AAG05910
ID AAG05910 standard; protein; 251 AA.
XX AC AAG05910;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2491.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Matches		7; Conservative	4; Mismatches	3; Indels	0; Gaps	0;
QY	1	CSLQEFLSHGGYVC	14			
		: : : :				
Db	28	CALKWFVNHGSTVC	41			
RESULT 29						
AAY59428						
ID	AAY59428 standard; protein; 313 AA.					
XX						
AC	AAY59428;					
XX						
DT	21-MAR-2000 (first entry)					
XX						
DE	Soybean inositol 1,3,4-triphosphate 5/6-kinase protein sequence.					
XX						
KW	Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;					
KW	myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;					
KW	animal feed.					
XX						
OS	Glycine max.					
XX						
PN	WO9955879-A1.					
XX						
PD	04-NOV-1999.					
XX						
PF	22-APR-1999; 99WO-US008790.					
XX						
PR	24-APR-1998; 98US-0082960P.					
XX						
PA	(DUPO) DU PONT DE NEMOURS & CO E I.					
XX						
PI	Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;					
XX						
DR	WPI; 2000-072179/06.					
DR	N-PSDB; AAZ48812.					
XX						
PT	Novel phytic acid biosynthetic enzymes used to alter the level of the					
PT	enzyme in transgenic plants.					
XX						
PS	Claim 5; Page 53-54; 63pp; English.					
XX						
CC	This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein					
CC	of the invention. ITK is a phytic acid (Myo-inositol 1,2,3,4,5,6-					
CC	hexaphosphate) biosynthetic enzyme. The ITK enzymes of the invention may					
CC	be prepared recombinantly and used to raise antibodies, which are used					
CC	for detecting the enzymes in situ in cells or in vitro in cell extracts.					
CC	The polynucleotides of the invention may be used to create transgenic					
CC	plants in which the ITK levels are present at higher or lower levels than					
CC	normal, or in cell types or developmental processes where they are not					
CC	normally found. This would alter the level of 1,3,4-triphosphate 5/6					
CC	kinase found in those cells. In addition, it may be desirable to					
CC	eliminate expression of ITK genes for certain applications. The					
CC	polynucleotides also serve as a source of probes and primers, which are					
CC	useful for genetic mapping, as markers for traits linked to those genes,					
CC	and to isolate homologous sequences from other species, as well as for					
CC	physical mapping. Decreased amounts of phytate and increased amounts of					
CC	available phosphate in animal feeds such as corn would lead to improved					
CC	feed efficiency. The proteins of the present invention lead to a better					
CC	understanding of the phytic acid biosynthesis pathway, allowing it to be					
CC	exploited for commercial uses, e.g. in animal feeds					
XX						
SQ	Sequence 313 AA;					

Query Match		55.6%;	Score 45;	DB 3;	Length 313;
Best Local Similarity		72.7%;	Pred. No. 59;		
Matches	8; Conservative	2;	Mismatches	1;	Indels 0; Gaps 0;
QY	3	LQEFLSHGGYV	13		
		:			
Db	177	LQEFVNHGGV	187		

RESULT 30	
AAG05909	
ID	AAG05909 standard; protein; 370 AA.
XX	
AC	AAG05909;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 2490.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 45; DB 3; Length 370;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
|:|:|:|:|:|:|
Db 147 CALKWFVNHGSTVC 160

RESULT 31
ABU27233
ID ABU27233 standard; protein; 536 AA.
XX
AC ABU27233;
XX

DT 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #12760.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Chlamydia trachomatis.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Walli D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
DR N-PSDB; ACA31103.
DR New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 55157; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 536 AA;
Query Match 55.6%; Score 45; DB 6; Length 536;
Best Local Similarity 63.6%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 2 SLQEFLSHGGY 12

Db |:|:|:|:|:|: 512 SIKEFLAHGGH 522
RESULT 32
AAU23148
ID AAU23148 standard; protein; 179 AA.
XX AC AAU23148;
XX DT 17-DEC-2001 (first entry)
XX DE Novel human enzyme polypeptide #234.
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX OS Homo sapiens.
XX WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001239.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
DR WPI; 2001-465566/50.
DR N-PSDB; AAS41018.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
PS Claim 11; SEQ ID NO 1144; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 179 AA;

Query Match 54.3%; Score 44; DB 4; Length 179;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGQYVC 14
| | | | | : | | | | |
Db 62 CPLPPMVSHGDFVC 75

RESULT 33
AAU87567
ID AAU87567 standard; protein; 179 AA.
XX
AC AAU87567;
XX
DT 05-JUN-2002 (first entry)
XX
DE Novel central nervous system protein #477.
XX

KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001332.
XX

PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 13-OCT-2000; 2000US-023937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-581633/65.
DR N-PSDB; ABK43897.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
PS Claim 9; SEQ ID NO 1085; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 54.3%; Score 44; DB 4; Length 179;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQFLSHGGYVC 14
Db 62 CPLPPMVSHGDFVC 75

RESULT 34
ADIS4882
ID ADIS4882 standard; protein; 179 AA.
XX
AC ADIS4882;
XX
DT 22-APR-2004 (first entry)
XX
DE Novel human protein seq id 1085.
XX
KW neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
KW antidiabetic; antirheumatic; antiarthritic; dermatological;
KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiant;
KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;

KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW amyotrophic lateral sclerosis; multiple sclerosis;
KW immune system disorder; diabetes; rheumatoid arthritis;
KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
KW inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; infectious disease;
KW HIV infection; hepatitis infection; bacterial infection;
KW fungal infection; parasitic infection; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
KW renal disorder; acute glomerulonephritis; pyelonephritis;
KW renal lithiasis; proliferative disorder; cancerous diseases; human.
XX
OS Homo sapiens.
XX
PN US2004018969-A1.
XX
PD 29-JAN-2004.
XX
PF 17-JAN-2001; 2001US-00764875.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2004-122079/12.
DR N-PSDB; ADI54284.
XX
PT New polypeptides and nucleic acid molecules, useful for detecting,
PT preventing, diagnosing, prognosticating, treating or ameliorating medical
PT conditions e.g. neural disorders, reproductive disorders or infectious
PT diseases.
XX
PS Claim 11; SEQ ID NO 1085; 413pp; English.
XX
CC The invention describes an isolated polypeptide comprising an amino acid
CC sequence at least 90% identical to: a polypeptide fragment, domain,
CC epitope, or full-length protein of any one of 607 amino acid sequences
CC (I) described in the specification; a polypeptide fragment of (I), or the
CC encoded sequence contained in (II), having biological activity; or a
CC variant, allelic variant, or a species homologue of (I). The polypeptides
CC and nucleic acid molecules are useful for detecting, preventing,
CC diagnosing, prognosticating, treating or ameliorating medical conditions
CC such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
CC immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic

Query Match 54.3%; Score 44; DB 8; Length 179;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLQEFLSHGGYVC 14
| | | : | | | : | |
Db 62 CPLPPMVSHGDFVC 75

RESULT 35
ABG26613
ID ABG26613 standard; protein; 184 AA.
XX
AC ABG26613;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26604.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX

PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS90800.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 56972; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 184 AA;

Query Match 54.3%; Score 44; DB 4; Length 184;
Best Local Similarity 50.0%; Pred. NO. 50;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLQEFLSHGYYVC 14
| | : | | : | |
Db 113 CPLPPMVSHGDFVC 126

RESULT 36
ABB03078
ID ABB03078 standard; protein; 187 AA.
XX
AC ABB03078;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human expressed polypeptide SEQ ID NO 51.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200155167-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001319.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.

PR 26-SEP-2000; 2000US-0235484P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465559/50.
DR N-PSDB; AAI99630.
XX
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with polypeptides and for
PT treating autoimmune diseases e.g., multiple sclerosis, rheumatoid
PT arthritis.
XX
PS Claim 11; SEQ ID NO 51; 504pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99614-AAI99654) and proteins
CC (ABB03062-ABB03085) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 187 AA;

Query Match 54.3%; Score 44; DB 4; Length 187;
Best Local Similarity 50.0%; Pred. NO. 51;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLQEFLSHGYYVC 14
| | : | | : | |
Db 62 CPLPPMVSHGDFVC 75

RESULT 37
AAU87270
ID AAU87270 standard; protein; 187 AA.
XX
AC AAU87270;
XX
DT 05-JUN-2002 (first entry)
XX
DE Novel central nervous system protein #180.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
OS Homo sapiens.
XX

XX WPI; 2001-581633/65.
DR N-PSDB; ABK43600.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
PS Claim 9; SEQ ID NO 788; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantaton, for supporting cell culture of
CC primary tissues, to regenerate tissues and in Chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 54.3%; Score 44; DB 4; Length 187;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
| | : | | : | |
Db 62 CPLPPMVSHGDFVC 75

RESULT 38
AD154585
ID AD154585 standard; protein; 187 AA.
XX
AC AD154585;
XX
DT 16-DEC-2004 (first entry)
XX
DE Novel human protein seq id 788.
XX
KW neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
KW antidiabetic; antirheumatic; antiarthritic; dermatological;
KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
KW cardiovascular-; antiarteriosclerotic; antiarrhythmic; cardiant;
KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW amyotrophic lateral sclerosis; multiple sclerosis;
KW immune system disorder; diabetes; rheumatoid arthritis;
KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
KW inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; infectious disease;
KW HIV infection; hepatitis infection; bacterial infection;
KW fungal infection; parasitic infection; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
KW renal disorder; acute glomerulonephritis; pyelonephritis;
KW renal lithiasis; proliferative disorder; cancerous diseases; human.

OS Homo sapiens.
XX
PN US2004018969-A1.
XX
PD 29-JAN-2004.
XX
PF 17-JAN-2001; 2001US-00764875.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2004-122079/12.
DR N-PSDB; ADI53987.
XX
PT New polypeptides and nucleic acid molecules, useful for detecting,
PT preventing, diagnosing, prognosticating, treating or ameliorating medical
PT conditions e.g. neural disorders, reproductive disorders or infectious
PT diseases.
XX
PS Claim 11; SEQ ID NO 788; 413pp; English.
XX
CC The invention describes an isolated polypeptide comprising an amino acid
CC sequence at least 90% identical to: a polypeptide fragment, domain,
CC epitope, or full-length protein of any one of 607 amino acid sequences
CC (I) described in the specification; a polypeptide fragment of (I), or the
CC encoded sequence contained in (II), having biological activity; or a
CC variant, allelic variant, or a species homologue of (I). The polypeptides
CC and nucleic acid molecules are useful for detecting, preventing,
CC diagnosing, prognosticating, treating or ameliorating medical conditions
CC such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
CC immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic

Query Match 54.3%; Score 44; DB 8; Length 187;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
| | : | | : | |
Db 62 CPLPPMVSHGDFVC 75

RESULT 39
AAG44384
ID AAG44384 standard; protein; 196 AA.
XX
AC AAG44384;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55591.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-013049P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 54.3%; Score 44; DB 3; Length 196;		
Best Local Similarity 63.6%; Pred. No. 53;		
Matches	7; Conservative	3; Mismatches 1; Indels 0; Gaps 0;
Qy	3 LQEFLSHGGYV 13	
	:	
Db	53 LQEFVNHGGVI 63	
RESULT 40		
AAG44383		
ID	AAG44383 standard; protein; 222 AA.	
XX		
AC	AAG44383;	
XX		
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55590.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	28-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.

PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 222;
Best Local Similarity 63.6%; Pred. No. 61;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGYYV 13
||||:||||:
Db 79 LQEFVNHGGVI 89

RESULT 41
ADJ32010
ID ADJ32010 standard; protein; 238 AA.

XX AC ADJ32010;

XX DT 15-APR-2004 (first entry)

XX DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #2.

XX KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.

XX OS Unidentified.

XX PN US2003229917-A1.

XX PD 11-DEC-2003.

XX PF 20-DEC-2002; 2002US-00324316.

XX PR 11-OCT-2000; 2000US-00686521.

XX PA (CAHO/) CAHOON R E.

XX PA (HITZ/) HITZ W D.

XX PA (PEAR/) PEARLSTEIN R W.

XX PA (CARL/) CARLSON T J.

XX PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;

XX DR WPI; 2004-052054/05.

XX PT New polynucleotide comprising a sequence encoding a polypeptide having

XX PT inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or

XX PT physical mapping for plant breeding.

XX PS Example 3; Fig 1; 32pp; English.

XX CC This invention relates to a novel isolated polynucleotide which comprises
CC a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
CC kinase activity and comprising a sequence having 80% identity based on
CC the Clustal alignment method with a fully defined sequence comprising 139
CC amino acids, or its complement. The polynucleotide is useful in genetic
CC or physical mapping for plant breeding. The present sequence is that of a
CC plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC exemplification of the invention.

XX SQ Sequence 238 AA;

Query Match 54.3%; Score 44; DB 8; Length 238;
Best Local Similarity 63.6%; Pred. No. 65;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
| | | | : | | | :
Db 133 LQEFVNHGGVI 143

RESULT 42
ADJ32009
ID ADJ32009 standard; protein; 238 AA.
XX
AC ADJ32009;
XX
DT 15-APR-2004 (first entry)
XX
DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #1.
XX
KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.
XX
OS Unidentified.
XX
PN US2003229917-A1.
XX
PD 11-DEC-2003.
XX
PF 20-DEC-2002; 2002US-00324316.
XX
PR 11-OCT-2000; 2000US-00686521.
XX
PA (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX
DR WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or physical mapping for plant breeding.
PT
PT
XX
PS Example 3; Fig 1; 32pp; English.
XX
CC This invention relates to a novel isolated polynucleotide which comprises a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-kinase activity and comprising a sequence having 80% identity based on the Clustal alignment method with a fully defined sequence comprising 139 amino acids, or its complement. The polynucleotide is useful in genetic or physical mapping for plant breeding. The present sequence is that of a plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the exemplification of the invention. Note: The SeqID numbers given in figure 1 do not correspond to those in the sequence listing for SeqIDs 2-16.
CC
CC
SQ Sequence 238 AA;

Query Match 54.3%; Score 44; DB 8; Length 238;
Best Local Similarity 63.6%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
| | | | : | | | :
Db 133 LQEFVNHGGVI 143

RESULT 43
ADJ32014
ID ADJ32014 standard; protein; 247 AA.
XX
AC ADJ32014;
XX
DT 15-APR-2004 (first entry)

XX Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #6.
DE
XX inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.
XX
OS Zea mays.
XX
PN US2003229917-A1.
XX
PD 11-DEC-2003.
XX
PF 20-DEC-2002; 2002US-00324316.
XX
PR 11-OCT-2000; 2000US-00686521.
XX
PA (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX
DR WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or physical mapping for plant breeding.
PT
PT
XX
PS Example 3; Fig 1; 32pp; English.
XX
CC This invention relates to a novel isolated polynucleotide which comprises a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-kinase activity and comprising a sequence having 80% identity based on the Clustal alignment method with a fully defined sequence comprising 139 amino acids, or its complement. The polynucleotide is useful in genetic or physical mapping for plant breeding. The present sequence is that of a plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the exemplification of the invention.
CC
CC
SQ Sequence 247 AA;

Query Match 54.3%; Score 44; DB 8; Length 247;
Best Local Similarity 63.6%; Pred. No. 68;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
| | | | : | | | :
Db 143 LQEFVNHGGVI 153

RESULT 44
ADJ32016
ID ADJ32016 standard; protein; 261 AA.
XX
AC ADJ32016;
XX
DT 15-APR-2004 (first entry)
XX
DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #8.
XX
KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.
XX
OS Oryza sativa.
XX
PN US2003229917-A1.
XX
PD 11-DEC-2003.
XX
PF 20-DEC-2002; 2002US-00324316.
XX
PR 11-OCT-2000; 2000US-00686521.

XX (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
PT inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
PT physical mapping for plant breeding.
XX
PS Example 3; Fig 1; 32pp; English.
XX
CC This invention relates to a novel isolated polynucleotide which comprises
CC a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
CC kinase activity and comprising a sequence having 80% identity based on
CC the Clustal alignment method with a fully defined sequence comprising 139
CC amino acids, or its complement. The polynucleotide is useful in genetic
CC or physical mapping for plant breeding. The present sequence is that of a
CC plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC exemplification of the invention.
XX
SQ Sequence 261 AA;

Query Match 54.3%; Score 44; DB 8; Length 261;
Best Local Similarity 63.6%; Pred. No. 72;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
Db 142 LQEFVNHGGVI 152
|||:||||:
142 LQEFVNHGGVI 152

RESULT 45
ADJ32018
ID ADJ32018 standard; protein; 279 AA.
XX
AC ADJ32018;
XX
DT 15-APR-2004 (first entry)
XX
DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #10.
XX
KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
KW physical mapping; plant breeding; plant; enzyme.
XX
OS Glycine max.
XX
PN US20032239917-A1.
XX
PD 11-DEC-2003.
XX
PF 20-DEC-2002; 2002US-00324316.
XX
PR 11-OCT-2000; 2000US-00686521.
XX
PA (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
PT inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
PT physical mapping for plant breeding.
XX
PS Example 3; Fig 1; 32pp; English.

XX This invention relates to a novel isolated polynucleotide which comprises
CC a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
CC kinase activity and comprising a sequence having 80% identity based on
CC the Clustal alignment method with a fully defined sequence comprising 139
CC amino acids, or its complement. The polynucleotide is useful in genetic
CC or physical mapping for plant breeding. The present sequence is that of a
CC plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC exemplification of the invention.
XX
SQ Sequence 279 AA;

Query Match 54.3%; Score 44; DB 8; Length 279;
Best Local Similarity 63.6%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
Db 173 LQEFVNHGGVI 183
|||:||||:
173 LQEFVNHGGVI 183

RESULT 46
AAG44382
ID AAG44382 standard; protein; 319 AA.
XX
AC AAG44382;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55589.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 319;
Best Local Similarity 63.6%; Pred. No. 88;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
Db 176 LQEFVNHGGVI 186

RESULT 47
AAY59421
ID AAY59421 standard; protein; 321 AA.
XX
AC AAY59421;
XX
DT 21-MAR-2000 (first entry)
XX
DE Corn inositol 1,3,4-triphosphate 5/6-kinase protein sequence.
XX
KW Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
KW myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
KW animal feed.
XX
OS Zea mays.
XX
PN WO9955879-A1.
XX
PD 04-NOV-1999.
XX
PF 22-APR-1999; 99WO-US008790.
XX
PR 24-APR-1998; 98US-0082960P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;
XX
DR WPI; 2000-072179/06.
DR N-PSDB; AAZ48805.
XX
PT Novel phytic acid biosynthetic enzymes used to alter the level of the
PT enzyme in transgenic plants.
XX
PS Claim 5; Page 44-45; 63pp; English.
XX
CC This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein
CC of the invention. ITK is a phytic acid (Myo-inositol 1,2,3,4,5,6-
CC hexaphosphate) biosynthetic enzyme. The ITK enzymes of the invention may
CC be prepared recombinantly and used to raise antibodies, which are used
CC for detecting the enzymes in situ in cells or in vitro in cell extracts.
CC The polynucleotides of the invention may be used to create transgenic
CC plants in which the ITK levels are present at higher or lower levels than
CC normal, or in cell types or developmental processes where they are not
CC normally found. This would alter the level of 1,3,4-triphosphate 5/6
CC kinase found in those cells. In addition, it may be desirable to
CC eliminate expression of ITK genes for certain applications. The
CC polynucleotides also serve as a source of probes and primers, which are
CC useful for genetic mapping, as markers for traits linked to those genes,
CC and to isolate homologous sequences from other species, as well as for
CC physical mapping. Decreased amounts of phytate and increased amounts of
CC available phosphate in animal feeds such as corn would lead to improved
CC feed efficiency. The proteins of the present invention lead to a better
CC understanding of the phytic acid biosynthesis pathway, allowing it to be
CC exploited for commercial uses, e.g. in animal feeds
XX
SQ Sequence 321 AA;

Query Match 54.3%; Score 44; DB 3; Length 321;
Best Local Similarity 63.6%; Pred. No. 88;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 63.6%; Pred. No. 89;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
Db 185 LQEFVNHGGVI 195

RESULT 48
AAY59423
ID AAY59423 standard; protein; 341 AA.
XX
AC AAY59423;
XX
DT 21-MAR-2000 (first entry)
XX
DE Rice inositol 1,3,4-triphosphate 5/6-kinase protein sequence.
XX
KW Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
KW myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
KW animal feed.
XX
OS Oryza sativa.
XX
PN WO9955879-A1.
XX
PD 04-NOV-1999.
XX
PF 22-APR-1999; 99WO-US008790.
XX
PR 24-APR-1998; 98US-0082960P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;
XX
DR WPI; 2000-072179/06.
DR N-PSDB; AAZ48807.
XX
PT Novel phytic acid biosynthetic enzymes used to alter the level of the
PT enzyme in transgenic plants.
XX
PS Claim 5; Page 47-48; 63pp; English.
XX
CC This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein
CC of the invention. ITK is a phytic acid (Myo-inositol 1,2,3,4,5,6-
CC hexaphosphate) biosynthetic enzyme. The ITK enzymes of the invention may
CC be prepared recombinantly and used to raise antibodies, which are used
CC for detecting the enzymes in situ in cells or in vitro in cell extracts.
CC The polynucleotides of the invention may be used to create transgenic
CC plants in which the ITK levels are present at higher or lower levels than
CC normal, or in cell types or developmental processes where they are not
CC normally found. This would alter the level of 1,3,4-triphosphate 5/6
CC kinase found in those cells. In addition, it may be desirable to
CC eliminate expression of ITK genes for certain applications. The
CC polynucleotides also serve as a source of probes and primers, which are
CC useful for genetic mapping, as markers for traits linked to those genes,
CC and to isolate homologous sequences from other species, as well as for
CC physical mapping. Decreased amounts of phytate and increased amounts of
CC available phosphate in animal feeds such as corn would lead to improved
CC feed efficiency. The proteins of the present invention lead to a better
CC understanding of the phytic acid biosynthesis pathway, allowing it to be
CC exploited for commercial uses, e.g. in animal feeds
XX
SQ Sequence 341 AA;

Query Match 54.3%; Score 44; DB 3; Length 341;
Best Local Similarity 63.6%; Pred. No. 94;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 49
ADC13500
ID ADC13500 standard; protein; 342 AA.
XX
AC ADC13500;
XX
DT 18-DEC-2003 (first entry)
XX
DE Protein which modulates inositol levels #4.
XX
KW inositol 1,3,4-triphosphate 5/6-kinases; ITPK; non-phytate phosphorous;
KW animal feed.
XX
OS Zea mays.
XX
PN WO2003027243-A2.
XX
PD 03-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030264.
XX
PR 27-SEP-2001; 2001US-0325308P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Wang H, Wu Y;
XX
DR WPI; 2003-371915/35.
DR N-PSDB; ADC13499.
XX
PT Novel nucleic acid molecule which modulates inositol 1,3,4-triphosphate
PT 5/6-kinases, useful to alter plant phenotype, to decrease phytate and
PT increase non-phytate phosphorus in corn or soy animal feedstuffs.
XX
PS Claim 1; SEQ ID NO 8; 88pp; English.
XX
CC The present invention relates to an isolated nucleic acid which modulates
CC the level of inositol 1,3,4-triphosphate 5/6-kinases (ITPK). The method
CC is useful for altering plant phenotype, by transforming a plant host
CC cell, and at least one polynucleotide of interest (POI), growing the
CC transformed host cell to modulate the activity of ITPK and the
CC polynucleotide of interest in the host cell and generating a transformed
CC plant with an altered phenotype. POI is myoinositol monophosphatase or
CC phytase, the activity of ITPK is down regulated, and the activity of the
CC polynucleotide of interest is up-regulated, or POI is inositol
CC polyphosphate kinase or myo-inositol 1-phosphate synthase and the
CC activity of ITPK and POI are each down-regulated. The method is useful
CC for modulating ITPK activity or levels in a host cell, preferably a plant
CC cell. This method reduces the level of phytate and increases the level of
CC non-phytate phosphorous. It is also useful for improving the nutritional
CC value of animal feed, by transforming a plant host cell to reduce phytate
CC content, growing the transformed host cell to modulate ITPK activity in
CC the host cell, generating a plant with the transformed genotype and
CC producing animal feed from the plant, where the animal feed has improved
CC nutritional value. The present sequence represents a protein which can
CC modulate the level of inositol.
XX
SQ Sequence 342 AA;

Query Match 54.3%; Score 44; DB 7; Length 342;
Best Local Similarity 63.6%; Pred. No. 94;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGQYV 13
||||:||||:
Db 186 LQEFVNHGGVI 196

RESULT 50
ADT59474
ID ADT59474 standard; protein; 352 AA.

ADT59474;
13-JAN-2005 (first entry)
Plant polypeptide, SEQ ID 9551.

Plant; transgenic; cold tolerance; growth rate; drought tolerance;
disease resistance; galactomannan production; plant growth regulator;
heat tolerance; herbicide tolerance; lignin production;
extreme osmotic condition tolerance; pathogens resistance;
pest resistance; yield improvement; seed oil yield; seed protein yield.

Viridiplantae.
US2004216190-A1.
28-OCT-2004.
18-DEC-2003; 2003US-00739930.
28-APR-2003; 2003US-00424599.
28-APR-2003; 2003US-00425115.
(KOVA/) KOVALIC D K.
Kovalic DK;
WPI; 2004-757369/74.

New recombinant DNA constructs useful in the field of biochemistry and
genetics, and in particular for producing transgenic plants with improved
biological characteristics.

Claim 2; SEQ ID NO 9551; 14pp; English.

The invention relates a recombinant DNA construct comprising a
polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
(SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
Arabidopsis, wheat and rape but the specification does not indicate which
sequences is derived from which organism. Also included is a method of
producing a plant having an improved property, comprising transforming a
plant with a recombinant DNA construct comprising a promoter region
functional in a plant cell operably joined to a polynucleotide encoding a
polypeptide associated with the property, and growing the transformed
plant. The property is selected from improving plant cold tolerance, for
manipulating growth rate in plant cells by modification of the cell cycle
pathway, for improving plant drought tolerance, for providing increased
resistance to plant disease, for galactomannan production, for production
of plant growth regulators, for improving plant heat tolerance, for
improving plant tolerance to herbicides, for increasing the rate of
homologous recombination in plants, for lignin production, for improving
plant tolerance to extreme osmotic conditions, for improving plant
tolerance to pathogens or pests, for yield improvement by modification of
photosynthesis, for modifying seed oil yield and/or content, for
modifying seed protein yield and/or content, for yield improvement by
modification of carbohydrate, nitrogen or phosphorus use and/or uptake
and for yield improvement by providing improved plant growth and
development under at least one stress condition. The polynucleotide may
also encode a plant transcription factor. The methods and compositions of
the present invention are useful in the field of biochemistry and
genetics, in particular for producing transgenic plants with improved
biological characteristics such as increased yield, improved nitrogen
flow, increasing plant tolerance to cold or heat, improving plant
tolerance to extreme osmotic and drought conditions, and improving plant
tolerance to plant pests or pathogens. They can also be used in physical
arrays of molecules, plant breeding markers, computer-based storage and
analysis systems. The present sequence is one of the 5544 plant protein
sequences of the invention. Note: The sequence data for this patent did
not form part of the printed specification, but was obtained in
electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX 50

Search completed: May 12, 2006, 10:37:20
Job time : 130.809 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16.9915 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-53
Perfect score: 81
Sequence: 1 CSLQEFLSHGGYVC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	48	59.3	660	2 A32378	probable ATP-depen
2	48	59.3	662	1 I84741	RNA helicase - mou
3	45	55.6	536	2 A81710	ribosomal protein
4	45	55.6	697	1 S13654	ATP-dependent RNA
5	44	54.3	319	2 JC5401	inositol 1,3,4-tri
6	44	54.3	429	2 A36220	transforming prote
7	42	51.9	293	2 T19954	hypothetical prote
8	42	51.9	338	2 T10544	inositol 1,3,4-tri
9	42	51.9	439	2 B72313	glutamine syntheta
10	41.5	51.2	113	2 H90876	probable minor tai
11	41	50.6	436	2 T23345	hypothetical prote
12	41	50.6	470	2 I38026	MLN 62 protein - h
13	40.5	50.0	880	2 G88925	protein T22H9.2 [i
14	40	49.4	392	2 PQ0446	DNA-binding protei
15	40	49.4	461	2 A46394	suppressor protein
16	40	49.4	497	2 A83884	L-arabinose isomer
17	40	49.4	569	2 A71556	probable S1 riboso
18	40	49.4	643	2 T27429	hypothetical prote
19	40	49.4	772	2 D56695	transducin-like en
20	39.5	48.8	113	2 D90854	minor tail protein
21	39.5	48.8	113	2 G90908	probable minor tai
22	39.5	48.8	113	2 B90823	probable minor tai
23	39	48.1	161	2 C95858	probable oxidoredu
24	39	48.1	187	2 B75576	oxidoreductase, ir
25	39	48.1	213	2 F71804	probable transcrip
26	39	48.1	525	2 T21527	hypothetical prote
27	39	48.1	532	2 C90154	primase SSO0140 [i
28	39	48.1	646	2 T35002	probable respirato
29	39	48.1	652	2 B83725	hypothetical prote

30	48.1	891	2	AG0265	alcohol dehydrogen
31	48.1	892	2	AE0650	alcohol dehydrogen
32	48.1	1049	1	S19421	ATP-dependent perm
33	48.1	1588	1	BVBYA1	3-dehydroquinat s
34	48.1	1642	2	T19130	hypothetical prote
35	47.5	1878	2	E86189	hypothetical prote
36	38.5	94	2	T03751	hypothetical prote
37	38	105	2	A82985	hypothetical prote
38	46.9	142	2	S46445	YMF46 protein - Ac
39	46.9	160	2	JE0136	lectin precursor -
40	46.9	213	2	E64690	response regulator
41	46.9	218	2	E82780	endonuclease III X
42	46.9	259	2	D75275	endonuclease III -
43	46.9	295	2	AH2340	transcription regu
44	46.9	310	2	T01266	starch synthase DU
45	46.9	323	2	D83708	phosphonates trans
46	46.9	421	2	E71363	hypothetical prote
47	46.9	425	2	E64657	probable glucose-6
48	46.9	447	1	A33504	glutamate dehydrog
49	46.9	447	2	AF0710	NADP-specific glut
50	46.9	471	2	E83130	conserved hypothet
51	46.9	670	2	AH2425	hypothetical prote
52	46.9	681	2	S76354	ABC1-type transpor
53	46.9	961	2	E86245	hypothetical prote
54	46.3	262	2	B82950	chromosome partiti
55	37.5	756	2	S47656	tMDC II protein -
56	37	101	2	C72537	hypothetical prote
57	37	177	2	S77911	hypothetical prote
58	37	210	2	E84943	DNA-(apurinic or a
59	37	214	2	D75576	oxidoreductase, ir
60	37	246	2	E82832	3-methyl-2-oxobuta
61	37	255	2	G87677	conserved hypothet
62	37	314	2	T28879	hypothetical prote
63	37	416	2	T26371	hypothetical prote
64	37	421	2	T23981	hypothetical prote
65	37	469	2	S23899	glutamate-ammonia
66	37	541	2	A48801	2-hydroxyacylsphin
67	37	716	2	T37710	hypothetical prote
68	37	833	2	S62136	CWH41 protein - ye
69	37	1026	2	T18220	chitin synthase (E
70	37	1142	2	T30272	hypothetical prote
71	37	1244	2	T23744	chromosome-associa
72	37	1620	2	T27283	hypothetical prote
73	37	2386	2	T39911	rad3 checkpoint pr
74	37	2531	2	S18188	notch protein homo
75	37	2818	2	B55282	neurofibromatosis-
76	37	2820	2	JC5196	neurofibromin I -
77	37	2825	2	I54352	neurofibromin - mo
78	36.5	379	2	T47507	probable protein k
79	36	156	1	VCTMHR	coat protein - tob
80	36	223	2	T44784	regulatory protein
81	36	230	2	S33161	Ig kappa chain - s
82	36	238	2	H84969	flagellar L-ring p
83	36	245	2	T51451	hypothetical prote
84	36	249	2	AF0735	probable bacteriop
85	36	267	2	A64375	thiamin biosynthes
86	36	271	2	JU0103	kanamycin kinase (
87	36	282	2	AF0332	hypothetical prote
88	36	287	2	JE0220	28k surface anti
89	36	335	2	I72789	interleukin 12 p40
90	36	351	2	F72020	leucine dehydrogen
91	36	351	2	E86605	leucine dehydrogen
92	36	360	2	T35783	probable secreted
93	36	364	2	T04994	hypothetical prote
94	36	397	2	B81839	probable periplasm
95	36	426	1	E70351	NADH2 dehydrogenas
96	36	438	2	T11903	NADH2 dehydrogenas
97	36	508	2	T03910	UDP-glucuronosyltr
98	36	522	2	T37536	cell polarity prot
99	36	531	2	T23835	hypothetical prote
100	36	535	2	T24309	hypothetical prote
101	36	555	2	AC0210	pectate disacchari
102	36	599	2	T16774	hypothetical prote

103	36	44.4	601	1	S11777	hydrogenase (EC 1.	176	35	43.2	740	2	B86638	GTP diphosphokinas
104	36	44.4	602	1	JQ0806	hydrogenase (EC 1.	177	35	43.2	860	1	QRHULD	LDL receptor precu
105	36	44.4	679	2	F83723	hypothetical prote	178	35	43.2	911	2	JC6016	chitin synthase (E
106	36	44.4	764	2	D84847	probable receptor-	179	35	43.2	916	2	JC2315	chitin synthase (E
107	36	44.4	780	2	D96803	probable beta-gala	180	35	43.2	916	2	A81089	DNA gyrase chain A
108	36	44.4	803	2	H87197	penicillin binding	181	35	43.2	916	2	B81853	DNA topoisomerase
109	36	44.4	834	2	T29821	hypothetical prote	182	35	43.2	916	2	S60779	DNA gyrase chain A
110	36	44.4	839	2	T45908	hypothetical prote	183	35	43.2	916	2	JT0396	reverse transcript
111	36	44.4	891	1	DEEC	hypothetical prote	184	35	43.2	943	2	JE0121	hypothetical 107.4
112	36	44.4	891	2	D85704	hypothetical prote	185	35	43.2	957	2	E86489	protein T32E20.19
113	36	44.4	891	2	E90846	acetaldehyde dehyd	186	35	43.2	961	2	AD0548	puative autotransp
114	36	44.4	928	1	VGBEBG	hypothetical prote	187	35	43.2	996	2	A71080	hypothetical prote
115	36	44.4	932	1	VGBEBC	acetaldehyde dehyd	188	35	43.2	1184	2	A96638	hypothetical prote
116	36	44.4	938	2	AB1530	glycoprotein GI pr	189	35	43.2	1206	2	D90085	hypothetical prote
117	36	44.4	938	2	AI1172	glycoprotein GI pr	190	35	43.2	1899	2	T49273	hypothetical prote
118	36	44.4	1056	2	H84777	transcription regu	191	35	43.2	1984	2	T13171	hypothetical prote
119	36	44.4	1059	2	T21891	probable kinesin-r	192	35	43.2	2204	1	RRNZNV	probable vitelloge
120	36	44.4	1156	2	T14899	hypothetical prote	193	35	43.2	2204	1	RRNZNV	genome polyprotein
121	35.5	43.8	334	2	T03764	aminophospholipid	194	34.5	42.6	4543	1	A53102	alpha-2-macroglobu
122	35.5	43.8	363	2	AC1700	protein RWD - rice	195	34.5	42.6	347	1	C46137	opsin, violet-sens
123	35.5	43.8	609	2	F84824	peptidoglycan synt	196	34.5	42.6	363	2	E83970	UDP-N-acetylglucos
124	35.5	43.8	682	2	T12968	hypothetical prote	197	34.5	42.6	524	2	S38539	disintegrin-like m
125	35.5	43.8	1097	1	S47220	hypothetical prote	198	34.5	42.6	670	2	I65967	disintegrin-like m
126	35.5	43.8	1148	2	T18770	protein kinase C (199	34.5	42.6	720	2	A36526	choline O-acetyltr
127	35.5	43.8	1412	2	T01610	probable calcium c	200	34.5	42.6	728	2	A24889	choline O-acetyltr
128	35	43.2	83	2	A82782	RNA-directed DNA p	201	34.5	42.6	915	1	A48225	subtilisin-like pr
129	35	43.2	103	2	AD1276	hypothetical prote	202	34.5	42.6	915	2	S34583	probable proprotei
130	35	43.2	103	2	AD1639	hypothetical prote	203	34.5	42.6	1548	2	F84316	serine proteinase
131	35	43.2	103	2	AD2089	hypothetical prote	204	34	42.0	40	2	F84316	hypothetical prote
132	35	43.2	174	2	G90432	hypothetical prote	205	34	42.0	68	2	H82124	heme exporter prot
133	35	43.2	187	2	A87512	conserved hypothet	206	34	42.0	98	2	AG2307	hypothetical prote
134	35	43.2	210	2	B64639	conserved hypothet	207	34	42.0	139	2	S67187	hypothetical prote
135	35	43.2	222	2	T30423	oxygen-insensitive	208	34	42.0	148	2	B82387	conserved hypothet
136	35	43.2	234	2	H85613	hypothetical prote	209	34	42.0	151	2	T32957	hypothetical prote
137	35	43.2	234	2	A36888	leucyl, phenylalan	210	34	42.0	152	2	F71973	hypothetical prote
138	35	43.2	234	2	B90750	leucyl, phenylalan	211	34	42.0	163	2	S43764	hypothetical prote
139	35	43.2	237	1	S56137	leucyl, phenylalan	212	34	42.0	163	2	S43764	mannose-binding le
140	35	43.2	249	2	T04436	membrane-bound tet	213	34	42.0	171	2	T49538	mannose-binding le
141	35	43.2	256	2	F70525	ankyrin 3 homolog	214	34	42.0	186	2	D64533	hypothetical prote
142	35	43.2	257	2	AB0353	probable glpQ2 pro	215	34	42.0	200	2	H81269	hypothetical prote
143	35	43.2	269	2	A48906	probable SpoU-fami	216	34	42.0	221	2	B81781	probable 3-isoprop
144	35	43.2	273	2	H71827	acetyltransferase	217	34	42.0	221	2	G81203	beta-phosphoglucom
145	35	43.2	290	2	T46470	nicotinate-nucleot	218	34	42.0	221	2	F81780	beta-phosphoglucom
146	35	43.2	303	2	A72365	hypothetical prote	219	34	42.0	226	2	F82091	probable anti-sigm
147	35	43.2	312	2	C84010	conserved hypothet	220	34	42.0	226	2	A70436	hypothetical prote
148	35	43.2	320	2	AC2315	3-oxoacyl-(acyl-ca	221	34	42.0	242	2	B86293	F7H2.20 protein -
149	35	43.2	330	2	AD2082	hypothetical prote	222	34	42.0	254	2	H72102	3-deoxy-manno-occu
150	35	43.2	344	2	F72406	iron(III) dicitrat	223	34	42.0	254	2	C86520	deoxyoctulonosis a
151	35	43.2	367	1	D70395	hypothetical prote	224	34	42.0	255	2	S74930	hypothetical prote
152	35	43.2	383	2	T34718	conserved hypothet	225	34	42.0	256	2	T15383	hypothetical prote
153	35	43.2	390	2	JC4023	probable integral	226	34	42.0	262	2	B90450	conserved hypothet
154	35	43.2	396	2	B83461	transforming growt	227	34	42.0	263	1	SNRTC2	proteasome endopep
155	35	43.2	400	2	C97147	hypothetical prote	228	34	42.0	263	2	A70163	hemolysin (tlyA) h
156	35	43.2	425	2	AE3358	probable glycosylt	229	34	42.0	264	1	WMBVT3	30K protein - toma
157	35	43.2	428	2	AF1984	phosphopyruvate hy	230	34	42.0	268	2	AG0332	conserved hypothet
158	35	43.2	430	2	H69364	hypothetical prote	231	34	42.0	269	1	JC1445	proteasome endopep
159	35	43.2	440	2	G87444	aspartyl-tRNA synt	232	34	42.0	271	1	F69656	naphthoate synthas
160	35	43.2	444	2	D71482	8-amino-7-oxononan	233	34	42.0	272	2	AI1283	dihydroxynaphthoic
161	35	43.2	448	1	Q4ADA7	probable thiophene	234	34	42.0	272	2	AD1655	dihydroxynaphthoic
162	35	43.2	449	1	Q4ADA2	maturation (pIva2)	235	34	42.0	274	2	F75328	conserved hypothet
163	35	43.2	449	1	Q4ADA5	maturation (pIva2)	236	34	42.0	275	2	B70640	hypothetical prote
164	35	43.2	452	2	S33932	maturation (pIva2)	237	34	42.0	275	2	G83107	probable two-compo
165	35	43.2	535	2	C95057	maturation protein	238	34	42.0	277	2	A97080	PTS system, fructo
166	35	43.2	535	2	F97926	CTP synthase [limp	239	34	42.0	288	2	B81650	endonuclease IV TC
167	35	43.2	544	2	S25101	CTP synthase (EC 6	240	34	42.0	288	2	B71491	probable endonucle
168	35	43.2	544	2	I55454	CTP synthase (EC 6	241	34	42.0	290	2	B71152	probable dTDP-4-de
169	35	43.2	548	2	T16642	neuroglycan C prec	242	34	42.0	290	2	AF1133	hypothetical prote
170	35	43.2	550	2	H64205	hypothetical prote	243	34	42.0	301	1	LNRT2	hepatic lectin 2 -
171	35	43.2	605	2	F84109	phosphomannomutase	244	34	42.0	310	2	S45130	autophagocytosis p
172	35	43.2	618	1	B43255	two-component sens	245	34	42.0	314	2	I38864	E1B 19K/Bcl-2-inte
173	35	43.2	626	2	T49625	hydrogenase (EC 1.	246	34	42.0	318	2	C81690	probable sodium-tr
174	35	43.2	626	2	S36364	glucan 1,4-alpha-g	247	34	42.0	325	1	A45470	hydroxymethylgluta
175	35	43.2	708	2	AH0192	glucan 1,4-alpha-g	248	34	42.0	325	1	A71887	probable GMP reduc
						alpha-galactosidas				327	1	F64626	probable GMP reduc

249	34	42.0	333	1	B53308	mosa protein - Rhi	322	33.5	41.4	2083	2	T42721	CRP-ductin-alpha p
250	34	42.0	335	2	C64380	porphobilinogen sy	323	33.5	41.4	2524	2	A35844	Xotch protein - Af
251	34	42.0	350	2	H96768	protein flavonol s	324	33	40.7	49	2	F82535	hypothetical prote
252	34	42.0	361	2	A87132	probable alcohol d	325	33	40.7	54	2	AH1965	hypothetical prote
253	34	42.0	361	2	G70862	probable adhE2 pro	326	33	40.7	95	2	H71256	conserved hypothet
254	34	42.0	364	2	B36313	hypothetical 42K p	327	33	40.7	98	2	T42909	hypothetical prote
255	34	42.0	370	2	S70157	cpSF protein, 40,6	328	33	40.7	104	2	F53275	Ig kappa-1 chain C
256	34	42.0	379	2	T16213	APX-1 protein homo	329	33	40.7	131	2	I56894	complement regulat
257	34	42.0	405	2	A38491	lariat-Gebranching	330	33	40.7	132	1	HAFGH	hemoglobin alpha-t
258	34	42.0	406	2	A54857	transcription fact	331	33	40.7	133	2	AB3465	hypothetical prote
259	34	42.0	406	2	A71965	hypothetical prote	332	33	40.7	147	2	A69933	hypothetical prote
260	34	42.0	420	2	T00629	hypothetical prote	333	33	40.7	154	2	A89962	6,7-dimethyl-8-rib
261	34	42.0	457	2	F92819	amine oxidase, fla	334	33	40.7	162	2	S43761	mannose-binding le
262	34	42.0	457	2	F97597	hypothetical prote	335	33	40.7	173	2	S77292	hypothetical prote
263	34	42.0	473	2	AC0479	glycerol-3-phospha	336	33	40.7	177	2	E87240	probable peptide m
264	34	42.0	473	2	T46999	hypothetical prote	337	33	40.7	177	2	E71327	hypothetical prote
265	34	42.0	473	2	AC0239	probable GntR-fami	338	33	40.7	193	2	AI0387	conserved hypothet
266	34	42.0	483	2	H85073	hypothetical prote	339	33	40.7	196	2	A99170	proteasome subunit
267	34	42.0	485	2	T01244	hypothetical prote	340	33	40.7	196	2	D87712	imidazoleglycerol-
268	34	42.0	488	2	T45964	LAX1 / AUX1-like p	341	33	40.7	208	2	F84942	riboflavin synthas
269	34	42.0	501	2	T31754	hypothetical prote	342	33	40.7	208	2	T41710	hypothetical prote
270	34	42.0	504	2	T34106	hypothetical prote	343	33	40.7	211	2	S77532	rehydrin - Synecho
271	34	42.0	504	2	C64398	hypothetical prote	344	33	40.7	214	2	G90582	adenylate kinase (
272	34	42.0	506	2	T33770	carboxypeptidase-1	345	33	40.7	214	2	A87078	probable 3-methylp
273	34	42.0	510	2	T48977	hypothetical prote	346	33	40.7	216	1	QQBEH8	9R protein - human
274	34	42.0	514	2	T25534	hypothetical prote	347	33	40.7	216	2	T44209	hypothetical prote
275	34	42.0	525	2	T24647	CTP synthase (pyrG	348	33	40.7	218	2	T44023	hypothetical prote
276	34	42.0	532	2	D69281	CTP synthase (UTP-	349	33	40.7	219	1	H64880	probable beta-phos
277	34	42.0	535	2	G97255	probable fructose-	350	33	40.7	219	2	A85753	probable beta-phos
278	34	42.0	553	2	G71543	steroid hormone re	351	33	40.7	219	2	H90865	hypothetical prote
279	34	42.0	560	2	S27874	CTP synthase (UTP-	352	33	40.7	235	2	E96811	amino acid ABC tra
280	34	42.0	578	2	D97555	CTP synthase [impo	353	33	40.7	239	2	A75293	PRE-MRNA SPLICING
281	34	42.0	578	2	AF2775	excinnuclease ABC c	354	33	40.7	243	2	T46221	hypothetical prote
282	34	42.0	594	1	E64622	excinnuclease ABC c	355	33	40.7	250	2	H84962	hypothetical prote
283	34	42.0	594	2	B71893	gamma-aminobutyric	356	33	40.7	252	2	C82507	hypothetical prote
284	34	42.0	627	2	B44409	probable acyl-CoA	357	33	40.7	253	2	T25768	hypothetical prote
285	34	42.0	627	2	JH0695	hypothetical prote	358	33	40.7	258	2	D71280	hypothetical prote
286	34	42.0	665	2	B96668	hypothetical prote	359	33	40.7	262	2	I49361	natural killer cel
287	34	42.0	693	2	T15152	hypothetical prote	360	33	40.7	262	2	A45813	T-cell surface gly
288	34	42.0	704	2	B84685	hypothetical prote	361	33	40.7	268	2	G90018	conserved hypothet
289	34	42.0	722	2	B71376	hypothetical prote	362	33	40.7	278	2	T38024	hypothetical prote
290	34	42.0	745	2	T23893	hypothetical prote	363	33	40.7	282	2	F82851	conserved hypothet
291	34	42.0	764	2	S48521	AKR1 protein - yea	364	33	40.7	290	2	AH3448	xdhc protein (assi
292	34	42.0	767	2	G96661	hypothetical prote	365	33	40.7	292	2	D71285	hypothetical prote
293	34	42.0	775	1	VPXRPG	outer layer protei	366	33	40.7	296	2	T15973	hypothetical prote
294	34	42.0	820	2	G88996	protein C17B7.5 [i	367	33	40.7	304	2	T12891	chemotaxis motA pr
295	34	42.0	833	2	AB0448	probable insectici	368	33	40.7	313	1	H69510	3-Hydroxy-3-methyl
296	34	42.0	841	2	T01011	hypothetical prote	369	33	40.7	320	2	E95315	probable D-isomer
297	34	42.0	972	2	A30363	glycoprotein GP330	370	33	40.7	322	2	AH1344	hypothetical prote
298	34	42.0	1009	2	T32464	hypothetical prote	371	33	40.7	322	2	AC1715	hypothetical prote
299	34	42.0	1184	2	T09484	cartilage intermed	372	33	40.7	323	2	AD3648	chemotaxis motA pr
300	34	42.0	1245	2	D71613	GAF domain protein	373	33	40.7	325	2	I57009	3-Hydroxy-3-methyl
301	34	42.0	1458	2	S36014	dynein heavy chain	374	33	40.7	326	2	AF0495	probable D-isomer
302	34	42.0	1463	2	C95032	hypothetical prote	375	33	40.7	327	2	T22407	hypothetical prote
303	34	42.0	1463	2	C97903	DNA-directed DNA p	376	33	40.7	334	2	F82038	adenosine deaminas
304	34	42.0	1650	2	S53457	dominant autoantig	377	33	40.7	337	2	B44478	probable cell grow
305	34	42.0	2531	2	A46019	notch-1 protein -	378	33	40.7	338	2	S55657	capsid protein 62
306	34	42.0	2555	2	A40043	notch protein homo	379	33	40.7	349	2	S22448	FMR1 protein - Pod
307	34	42.0	2670	2	A46719	inositol 1,4,5-tri	380	33	40.7	363	2	I38937	DNA/RNA-binding pr
308	34	42.0	2671	2	A49873	inositol 1,4,5-tri	381	33	40.7	365	2	A34895	5S RNA-binding pro
309	34	42.0	3788	2	T30851	lysosomal traffick	382	33	40.7	367	2	S47649	lignin peroxidase
310	34	42.0	3788	2	T13960	beige protein homo	383	33	40.7	368	2	F98282	histidinol-phospha
311	34	42.0	3796	2	T18514	lysosomal traffick	384	33	40.7	368	2	AB3001	histidinol-phospha
312	34	42.0	4544	1	S02392	alpha-2-macroglobu	385	33	40.7	373	2	A44478	probable cell grow
313	34	42.0	4660	2	T42737	gp330 protein prec	386	33	40.7	377	1	A48711	protein-tyrosine-p
314	34	41.4	226	2	A46477	membrane-bound imm	387	33	40.7	378	2	D69793	conserved hypothet
315	33.5	41.4	307	2	C70734	hypothetical prote	388	33	40.7	379	2	E81081	hypothetical prote
316	33.5	41.4	363	2	JC1275	phospho-N-acetylm	389	33	40.7	391	2	F81861	hypothetical prote
317	33.5	41.4	376	2	T16116	hypothetical prote	390	33	40.7	403	2	C64998	hypothetical prote
318	33.5	41.4	429	2	F84015	maltose/maltodextr	391	33	40.7	406	2	S42394	G-box-binding prot
319	33.5	41.4	438	2	C86244	DnaJ homolog, 4706	392	33	40.7	407	2	F85867	probable sulfatase
320	33.5	41.4	849	2	H84668	Mutator-like trans	393	33	40.7	407	2	E91023	probable sulfatase
321	33.5	41.4	1510	2	C84727	probable glucan sy	394	33	40.7	416	2	D84638	probable sarcosine

395 33 40.7 425 2 H70793 hypothetical prote
396 33 40.7 427 2 F75406 histidyl-trNA synt
397 33 40.7 431 2 AB0410 phosphopyruvate hy
398 33 40.7 432 1 NOEC phosphopyruvate hy
399 33 40.7 432 2 H85928 enolase [imported]
400 33 40.7 432 2 G91083 enolase [imported]
401 33 40.7 432 2 AC0859 enolase [imported]
402 33 40.7 432 2 G90268 conserved hypothet
403 33 40.7 433 2 B81992 NADH2 dehydrogenas
404 33 40.7 433 2 D81222 NADH dehydrogenase
405 33 40.7 433 2 C84619 probable serine ca
406 33 40.7 439 1 ISHIX xylose isomerase (
407 33 40.7 442 2 C81370 probable thiophene
408 33 40.7 456 2 E69391 hypothetical prote
409 33 40.7 459 2 T51558 probable flavonol
410 33 40.7 460 2 T04579 hypothetical prote
411 33 40.7 464 2 T47710 glucuronosyl trans
412 33 40.7 465 2 D96567 F6D8.13 [imported]
413 33 40.7 468 2 H70201 UDP-N-acetylmurama
414 33 40.7 469 2 B37837 probable alpha-amy
415 33 40.7 472 2 T51559 probable flavonol
416 33 40.7 479 2 B69764 transcription regu
417 33 40.7 485 2 S36772 E-selectin - bovin
418 33 40.7 493 2 S32037 finger protein XFG
419 33 40.7 497 2 T27827 hypothetical prote
420 33 40.7 506 2 C81704 monooxygenase-rela
421 33 40.7 512 2 T40576 probable structure
422 33 40.7 519 2 JC4762 RNA-directed RNA p
423 33 40.7 526 2 T25535 hypothetical prote
424 33 40.7 527 2 I84483 protein-tyrosine k
425 33 40.7 533 1 S75536 NADH2 dehydrogenas
426 33 40.7 540 2 T15573 hypothetical prote
427 33 40.7 541 2 JC5423 2-hydroxyacylsphin
428 33 40.7 550 1 A34576 crystal protein pr
429 33 40.7 553 2 C81698 pyrophosphate-fruc
430 33 40.7 567 2 T25771 hypothetical prote
431 33 40.7 596 1 S11969 hydrogenase (EC 1.
432 33 40.7 598 2 T42592 protein-serine/thr
433 33 40.7 599 2 B72368 conserved hypothet
434 33 40.7 610 2 A35046 E-selectin precurs
435 33 40.7 628 2 A87596 hypothetical prote
436 33 40.7 633 2 C70358 hydrogenase (EC 1.
437 33 40.7 634 2 S33339 somatotropin recep
438 33 40.7 635 1 T37835 probable phosphoes
439 33 40.7 638 2 A33991 somatotropin recep
440 33 40.7 638 2 S12136 somatotropin recep
441 33 40.7 638 2 B28176 somatotropin recep
442 33 40.7 640 2 S62747 homeotic protein A
443 33 40.7 642 2 E69144 probable formate C
444 33 40.7 653 2 T51895 hypothetical prote
445 33 40.7 662 2 G87630 methylnalonyl-CoA
446 33 40.7 665 2 T18979 hypothetical prote
447 33 40.7 712 1 A47331 ribonucleoside-tri
448 33 40.7 712 2 AC1058 ribonucleoside-tri
449 33 40.7 712 2 AG0419 ribonucleoside-tri
450 33 40.7 712 2 G91280 anaerobic ribonucl
451 33 40.7 712 2 G86121 anaerobic ribonucl
452 33 40.7 721 2 T08956 AIM1 protein - Ara
453 33 40.7 739 2 A83015 primosomal protein
454 33 40.7 775 2 AH2803 phosphoribosylform
455 33 40.7 775 2 H97582 phosphoribosylform
456 33 40.7 790 1 S77032 ABC transporter sl
457 33 40.7 830 2 A84951 DNA topoisomerase
458 33 40.7 868 1 VGBE31 glycoprotein B - h
459 33 40.7 872 2 S73785 DNA polymerase III
460 33 40.7 879 1 QRRULD LDL receptor precu
461 33 40.7 923 2 H83248 DNA gyrase subunit
462 33 40.7 942 2 C83861 ATP-dependent DNA
463 33 40.7 946 1 A47299 ror-related recept
464 33 40.7 953 2 T40643 probable serine th
465 33 40.7 1020 2 D86402 protein envelope C
466 33 40.7 1020 2 T51925 Ca2+-transporting
467 33 40.7 1020 2 T51926 Ca2+-transporting

468 33 40.7 1026 2 A49750 beta-galactosidase
469 33 40.7 1054 2 T14189 hypothetical prote
470 33 40.7 1133 1 EGRT epidermal growth f
471 33 40.7 1133 2 T12529 hypothetical prote
472 33 40.7 1207 1 EGHU epidermal growth f
473 33 40.7 1217 1 EGMSMG epidermal growth f
474 33 40.7 1274 2 I40813 neurotoxin type F
475 33 40.7 1304 2 T14073 dynein lb heavy ch
476 33 40.7 1346 2 T14849 cytoadherence link
477 33 40.7 1449 2 T20181 hypothetical prote
478 33 40.7 1514 2 S70099 hypothetical prote
479 33 40.7 1518 2 PQ0221 polyprotein - plum
480 33 40.7 1573 2 AF3514 glutamate synthase
481 33 40.7 1583 2 F86366 protein F26F24.8 [
482 33 40.7 1584 2 T20180 hypothetical prote
483 33 40.7 1586 2 T20179 hypothetical prote
484 33 40.7 1597 2 S65053 genome polyprotein
485 33 40.7 1699 2 T14074 complement compone
486 33 40.7 2264 1 GNVVTB genome polyprotein
487 33 40.7 2276 2 T00076 hypothetical prote
488 33 40.7 2330 1 RRIWMV genome polyprotein
489 33 40.7 2962 2 T19756 hypothetical prote
490 33 40.7 3131 2 S39842 enniatin synthetas
491 33 40.7 3225 2 D81702 adherence factor T
492 33 40.7 3635 2 T10053 laminin alpha 5 ch
493 33 40.7 3973 2 B71612 hypothetical prote
494 32.5 40.1 153 2 PN0103 hypothetical 17K p
495 32.5 40.1 196 2 AF3358 methyltransferase
496 32.5 40.1 258 2 AB1956 hypothetical prote
497 32.5 40.1 316 2 S57839 CPC2 protein - Neu
498 32.5 40.1 363 2 AC1329 peptidoglycan synt
499 32.5 40.1 409 2 B37753 NadR protein - Sal
500 32.5 40.1 476 2 H83265 probable D-alanyl-
501 32.5 40.1 520 2 B71443 hypothetical prote
502 32.5 40.1 2109 2 I38414 transcription fact
503 32.5 40.1 2946 2 T15840 hypothetical prote
504 32 39.5 53 2 AG0014 hypothetical prote
505 32 39.5 59 2 G81250 probable transcrit
506 32 39.5 63 2 T41452 probable protein t
507 32 39.5 65 2 C82542 hypothetical prote
508 32 39.5 80 2 AH2958 hypothetical prote
509 32 39.5 80 2 F98324 hypothetical prote
510 32 39.5 87 2 T03315 gene l8 protein -
511 32 39.5 91 2 JE0321 ribosomal protein
512 32 39.5 91 2 C83414 hypothetical prote
513 32 39.5 102 2 T32603 hypothetical prote
514 32 39.5 116 2 D81945 probable pilus bio
515 32 39.5 116 2 G81159 probable type IV p
516 32 39.5 122 2 B23843 serum amyloid prot
517 32 39.5 122 2 A23843 serum amyloid prot
518 32 39.5 122 2 I49496 amyloid A - mouse
519 32 39.5 127 2 F83803 hypothetical prote
520 32 39.5 132 2 T26527 hypothetical prote
521 32 39.5 140 2 AG2146 two-component resp
522 32 39.5 140 2 AC1264 hypothetical prote
523 32 39.5 140 2 AE1626 hypothetical prote
524 32 39.5 149 2 AB1395 hypothetical prote
525 32 39.5 149 2 AE1770 hypothetical prote
526 32 39.5 157 2 S19735 lectin precursor -
527 32 39.5 157 2 T40106 conserved hypothet
528 32 39.5 157 2 F87509 conserved hypothet
529 32 39.5 163 2 B84731 hypothetical prote
530 32 39.5 163 2 T10220 hypothetical prote
531 32 39.5 167 2 E83743 4-hydroxybenzoyl-C
532 32 39.5 173 2 A70017 4-hydroxybenzoyl-C
533 32 39.5 180 2 F82187 TonB system transp
534 32 39.5 183 1 C70008 pyrazinamidase/nic
535 32 39.5 184 2 S77751 probable ABC-type
536 32 39.5 190 2 F87412 conserved hypothet
537 32 39.5 192 2 H82467 hypothetical prote
538 32 39.5 193 2 D81006 conserved hypothet
539 32 39.5 195 2 E27733 nifQ protein - Azo
540 32 39.5 196 2 D69275 conserved hypothet

541	32	39.5	200	2	I48615	gene KIS protein -
542	32	39.5	203	2	G70501	probable 3-methylp
543	32	39.5	204	2	B75539	probable phosphogl
544	32	39.5	212	2	H83208	endonuclease III p
545	32	39.5	218	2	B47712	myelin/oligodendro
546	32	39.5	218	2	A64201	DNA primase (dnaE)
547	32	39.5	220	2	H87439	ParA family protei
548	32	39.5	220	2	S75751	hypothetical prote
549	32	39.5	221	2	B83868	DNA-3-methyladenin
550	32	39.5	223	2	H82482	transcription regu
551	32	39.5	225	2	C85475	hypothetical prote
552	32	39.5	230	2	C97104	hypothetical prote
553	32	39.5	232	2	T15146	hypothetical prote
554	32	39.5	233	2	S75267	hypothetical prote
555	32	39.5	235	2	A10365	probable two-compo
556	32	39.5	240	1	QRECGQ	glutamine transpor
557	32	39.5	240	2	A10600	glutamine transpor
558	32	39.5	240	2	G90739	hypothetical prote
559	32	39.5	240	2	A85590	hypothetical prote
560	32	39.5	241	2	A87712	endonuclease III [
561	32	39.5	241	2	F81997	leucyl/phenylalan
562	32	39.5	247	2	A55717	myelin/oligodendro
563	32	39.5	248	2	AH3474	DNA- (apurinic or a
564	32	39.5	248	2	T33281	hypothetical prote
565	32	39.5	250	2	AE0621	probable prophage
566	32	39.5	252	2	JC5237	osmotin-like prote
567	32	39.5	257	2	AD1996	3-methyl-2-oxobuta
568	32	39.5	260	2	D75404	phosphatidylglycer
569	32	39.5	262	2	AF2698	NAD/NADP dependent
570	32	39.5	265	2	AH3044	conserved hypothet
571	32	39.5	265	2	D98241	hypothetical prote
572	32	39.5	265	2	S74449	hypothetical prote
573	32	39.5	268	2	H87474	conserved hypothet
574	32	39.5	273	2	D89873	naphthoate synthas
575	32	39.5	275	2	E82554	extragenic suppress
576	32	39.5	278	2	F86651	LPS biosynthesis p
577	32	39.5	278	2	H87663	hypothetical prote
578	32	39.5	280	2	H95901	conserved hypothet
579	32	39.5	290	2	E84797	hypothetical prote
580	32	39.5	293	2	F97480	probable dehydroge
581	32	39.5	293	2	I39913	NADH2 dehydrogenas
582	32	39.5	300	2	T20905	hypothetical prote
583	32	39.5	303	2	H69551	acyl carrier prote
584	32	39.5	306	2	I49068	protein kinase Sry
585	32	39.5	307	2	AF2402	hypothetical prote
586	32	39.5	309	2	E83243	probable transcrip
587	32	39.5	312	2	AC3077	flavoprotein oxido
588	32	39.5	312	2	F98209	hypothetical prote
589	32	39.5	312	2	S50430	hypothetical prote
590	32	39.5	317	2	S63358	hypothetical prote
591	32	39.5	319	2	A12125	hypothetical prote
592	32	39.5	321	2	T01477	protein kinase hom
593	32	39.5	321	2	AE1282	conserved hypothet
594	32	39.5	321	2	A11653	conserved hypothet
595	32	39.5	323	2	B97457	BH3305 conserved h
596	32	39.5	323	2	AD2675	conserved hypothet
597	32	39.5	323	2	D86216	protein T23G18.8 [
598	32	39.5	324	2	B84064	glycerate dehydrog
599	32	39.5	326	2	G98231	exoO protein [limpo
600	32	39.5	326	2	AF3054	succinoglycan bios
601	32	39.5	326	2	T10166	restriction endonu
602	32	39.5	326	2	S74844	lmbp protein - Syn
603	32	39.5	328	2	AE0301	conserved hypothet
604	32	39.5	332	2	A86892	glycosyltransferas
605	32	39.5	342	2	A48258	dopamine receptor
606	32	39.5	346	2	A96926	endoglucanase, ami
607	32	39.5	348	1	Z1BPF1	gene I protein - p
608	32	39.5	348	1	Z1BPF1	gene I protein - p
609	32	39.5	348	1	Z1BPM3	gene I protein - p
610	32	39.5	348	2	T16076	LIM protein - Caen
611	32	39.5	348	2	G86270	hypothetical prote
612	32	39.5	353	2	AB2483	hypothetical prote
613	32	39.5	354	2	AC0389	phosphate binding
1	32	39.5	357	2	T05625	cinnamyl-alcohol d
2	32	39.5	358	2	F98289	lps biosynthesis r
3	32	39.5	359	2	F90055	conserved hypothet
4	32	39.5	362	2	S33733	G protein-coupled
5	32	39.5	370	2	G90248	conserved hypothet
6	32	39.5	373	2	JC4737	G protein-coupled
7	32	39.5	373	2	JC4162	p2y receptor - bov
8	32	39.5	374	2	E95361	probable muconate
9	32	39.5	381	1	B69030	conserved hypothet
10	32	39.5	381	2	D64525	GDP-D-mannose dehy
11	32	39.5	382	2	F96009	hypothetical membr
12	32	39.5	383	2	AH0139	galactokinase (EC
13	32	39.5	383	2	T04641	homeobox protein 1
14	32	39.5	390	2	A70656	hypothetical prote
15	32	39.5	392	1	C69851	macrolide glycosyl
16	32	39.5	393	2	AB2752	enolase [imported]
17	32	39.5	394	2	C96728	hypothetical prote
18	32	39.5	397	2	H84277	orc / cell divisio
19	32	39.5	399	2	T06251	GTP binding protei
20	32	39.5	400	2	G01977	d3 dopamine recept
21	32	39.5	404	2	AD2994	glycosyltransferas
22	32	39.5	406	2	T24301	hypothetical prote
23	32	39.5	408	2	E75452	hypothetical prote
24	32	39.5	408	2	G75376	probable transposa
25	32	39.5	410	2	S74705	probable transposa
26	32	39.5	412	2	H72741	hypothetical prote
27	32	39.5	413	2	C81278	probable thiazole
28	32	39.5	414	2	B98354	hypothetical prote
29	32	39.5	416	1	S33777	hypothetical prote
30	32	39.5	424	2	A97533	hepsin (EC 3.4.21.
31	32	39.5	425	2	T38548	enolase (2-phospho
32	32	39.5	425	2	AC2756	hypothetical zinc-
33	32	39.5	433	2	A84619	MFS permease limpo
34	32	39.5	435	2	E84618	probable serine ca
35	32	39.5	437	2	D84619	probable serine ca
36	32	39.5	439	2	D71115	hypothetical prote
37	32	39.5	444	2	T15310	hypothetical prote
38	32	39.5	449	2	E72383	dihydrolipoamide d
39	32	39.5	449	2	T16891	protein phosphatas
40	32	39.5	454	2	A38643	protein kinase (EC
41	32	39.5	455	2	S67775	hypothetical prote
42	32	39.5	459	2	S59731	mRNA guanylyltrans
43	32	39.5	467	2	T02238	glucosyl transfera
44	32	39.5	468	2	AG2899	dihydrolipoamide d
45	32	39.5	468	2	H97674	dihydrolipoamide d
46	32	39.5	468	2	B95338	DegP4 protease lik
47	32	39.5	480	2	T46925	hypothetical prote
48	32	39.5	481	2	S62427	G-protein signalin
49	32	39.5	483	2	A39676	protein kinase Sry
50	32	39.5	484	2	S53641	protein kinase clk
51	32	39.5	485	2	T39508	conserved hypothet
52	32	39.5	486	2	S10772	2-hydroxymuconic s
53	32	39.5	486	2	H75060	hydrogenase (EC 1.
54	32	39.5	486	2	S30959	gene 14 protein -
55	32	39.5	488	2	T06260	NADH2 dehydrogenas
56	32	39.5	489	1	VGXPMV	surface glycoprote
57	32	39.5	491	2	T52398	hypothetical prote
58	32	39.5	493	2	T33569	hypothetical prote
59	32	39.5	498	2	B97537	probable transmemb
60	32	39.5	504	2	A38221	chitinase (EC 3.2.
61	32	39.5	522	2	S28743	NADH2 dehydrogenas
62	32	39.5	524	2	F70579	probable cell divi
63	32	39.5	531	2	T15329	hypothetical prote
64	32	39.5	535	2	C86686	Ctp synthetase [im
65	32	39.5	542	2	T28719	hypothetical prote
66	32	39.5	546	2	JC4798	seizure-related me
67	32	39.5	562	2	H69545	probable fatty-aci
68	32	39.5	564	2	T49322	related to RNA-bin
69	32	39.5	567	2	AG2008	hypothetical prote
70	32	39.5	570	2	T46911	hypothetical prote
71	32	39.5	580	2	H87963	protein Y54E2A.1 [
72	32	39.5	589	2	E97376	sulfite reductase
73	32	39.5	589	2	AD2594	hypothetical prote

687 32 39.5 594 1 WZBEE2 protein-serine/thr
688 32 39.5 596 2 A85438 hypothetical prote
689 32 39.5 597 2 AE0721 hydrogenase-1 larg
690 32 39.5 598 2 I51368 gamma-aminobutyric
691 32 39.5 598 2 F46027 gamma-aminobutyric
692 32 39.5 599 1 ACRTGT gamma-aminobutyric
693 32 39.5 599 2 S11073 gamma-aminobutyric
694 32 39.5 601 2 D83583 probable acyl-CoA
695 32 39.5 602 2 H86468 protein F12K21.20
696 32 39.5 617 2 D90487 maltose ABC transp
697 32 39.5 618 2 S09251 hydrogenase (EC 1.
698 32 39.5 619 1 JH0776 hydrogenase (EC 1.
699 32 39.5 620 1 S33253 protein-tyrosine k
700 32 39.5 621 2 I38467 low density lipopr
701 32 39.5 625 1 A43030 protein-tyrosine k
702 32 39.5 640 2 S49932 MET30 protein - ye
703 32 39.5 642 2 G69786 ABC transporter (A
704 32 39.5 643 2 T19135 cholecystokinin ty
705 32 39.5 661 2 F96665 protein F22C12.14
706 32 39.5 668 2 T15305 hypothetical prote
707 32 39.5 675 2 T04644 hypothetical prote
708 32 39.5 699 2 T14904 NADPH-ferrihemopro
709 32 39.5 700 2 A84243 phosphoribosylform
710 32 39.5 700 2 A96690 hypothetical prote
711 32 39.5 703 2 AG0242 probable membrane
712 32 39.5 705 2 F70475 VacB protein (ribo
713 32 39.5 706 2 D82452 anaerobic ribonucl
714 32 39.5 713 2 T21201 hypothetical prote
715 32 39.5 719 2 A30047 enhancer of split
716 32 39.5 721 2 C71014 hypothetical prote
717 32 39.5 733 2 A95891 probable N-methylh
718 32 39.5 754 2 S50601 hypothetical prote
719 32 39.5 759 2 T41295 protein transport
720 32 39.5 763 2 E96693 probable terpene s
721 32 39.5 771 2 T01315 hypothetical prote
722 32 39.5 771 2 S35681 ESG protein - mous
723 32 39.5 784 2 T18452 hypothetical prote
724 32 39.5 788 1 B4AG58 virB4 protein prec
725 32 39.5 789 1 B4AGA6 virB4 protein prec
726 32 39.5 789 2 AH3248 component of type
727 32 39.5 796 2 AG1849 serine/threonine k
728 32 39.5 797 2 E84642 hypothetical prote
729 32 39.5 829 2 T32744 hypothetical prote
730 32 39.5 835 2 T26086 hypothetical prote
731 32 39.5 848 2 B89042 protein F14F9.3 [i
732 32 39.5 857 2 B84182 hypothetical prote
733 32 39.5 858 2 JC7683 taste receptor T1R
734 32 39.5 882 2 S41034 hypothetical prote
735 32 39.5 888 2 S64016 probable regulator
736 32 39.5 901 2 G71286 probable pyruvate,
737 32 39.5 951 2 T08987 probable cadmium-t
738 32 39.5 965 2 S62935 hypothetical prote
739 32 39.5 977 2 H84469 hypothetical prote
740 32 39.5 985 1 DJBEI1 DNA-directed DNA p
741 32 39.5 1034 2 JC5569 serine-threonine k
742 32 39.5 1050 2 JW0092 hypothetical prote
743 32 39.5 1052 2 T04439 hypothetical prote
744 32 39.5 1086 2 T24214 hypothetical prote
745 32 39.5 1086 2 T40354 hypothetical prote
746 32 39.5 1127 2 G71274 atrial natriuretic
747 32 39.5 1137 2 T18625 hypothetical prote
748 32 39.5 1147 2 T35781 hypothetical prote
749 32 39.5 1189 2 T51491 hypothetical prote
750 32 39.5 1191 2 T13850 gene u-shaped prot
751 32 39.5 1246 2 T51085 related to protein
752 32 39.5 1251 2 JH0256 botulinum neurotox
753 32 39.5 1252 2 S21178 botulinum neurotox
754 32 39.5 1314 2 G02870 KIAA0197 protein -
755 32 39.5 1318 2 T39066 hypothetical prote
756 32 39.5 1326 2 H89134 protein F25G6.9 [i
757 32 39.5 1407 2 S59823 probable membrane
758 32 39.5 1420 2 T02644 ABC-type transport
759 32 39.5 1449 2 B84426 hypothetical prote

760 32 39.5 1459 2 T24088 hypothetical prote
761 32 39.5 1520 1 TVFFA protein-tyrosine k
762 32 39.5 1545 2 F98262 hypothetical prote
763 32 39.5 1581 2 AD3022 glutamate synthase
764 32 39.5 1627 2 AE2109 two-component hybr
765 32 39.5 1722 2 A88470 protein C28H8.3 [i
766 32 39.5 1944 2 A55117 tag24 protein - mo
767 32 39.5 1957 2 S68453 sodium channel pro
768 32 39.5 2194 1 GNNYE7 genome polyprotein
769 32 39.5 2318 2 S45306 notch 3 protein -
770 32 39.5 2321 2 S78549 notch3 protein - h
771 32 39.5 3507 2 T34513 hypothetical prote
772 32 39.5 3759 2 A35085 trithorax protein
773 32 39.5 4391 2 A38096 perlecan precursor
774 31.5 38.9 113 2 T44561 hypothetical prote
775 31.5 38.9 173 2 AD3517 probable carbonyl
776 31.5 38.9 236 2 T33589 hypothetical prote
777 31.5 38.9 290 2 T02300 GTP-binding regula
778 31.5 38.9 325 2 T06784 GTP-binding protei
779 31.5 38.9 325 2 T09613 probable GTP-bind
780 31.5 38.9 328 2 T16970 GTP-binding protei
781 31.5 38.9 349 2 C72097 fructose-bisphosph
782 31.5 38.9 349 2 A86526 1,6-fructose biph
783 31.5 38.9 450 2 D86356 hypothetical prote
784 31.5 38.9 670 2 G96790 hypothetical prote
785 31.5 38.9 695 2 S49163 transferrin precur
786 31.5 38.9 792 2 D95351 VirB4 type IV secr
787 31.5 38.9 801 1 S00943 glucose dehydrogen
788 31.5 38.9 881 2 F69438 conserved hypothet
789 31.5 38.9 1088 2 H84604 probable cellulose
790 31.5 38.9 1275 2 T41523 hypothetical rho1
791 31.5 38.9 1599 2 T16210 hypothetical prote
792 31 38.3 18 2 S33645 hypothetical prote
793 31 38.3 61 2 E90761 hypothetical prote
794 31 38.3 63 2 D97714 hypothetical prote
795 31 38.3 64 2 T21841 hypothetical prote
796 31 38.3 69 1 D64989 hypothetical prote
797 31 38.3 69 2 G91014 heme exporter prot
798 31 38.3 69 2 A85859 heme exporter prot
799 31 38.3 80 2 AC2394 heme exporter prot
800 31 38.3 85 2 A85625 hypothetical prote
801 31 38.3 89 2 G64381 ribosomal protein
802 31 38.3 102 2 S49794 hypothetical prote
803 31 38.3 108 2 T16555 hypothetical prote
804 31 38.3 112 2 C71221 phospholipase A2 h
805 31 38.3 121 1 PC4024 hypothetical prote
806 31 38.3 123 2 D96496 hypothetical prote
807 31 38.3 128 2 AC3164 hypothetical prote
808 31 38.3 131 2 T28801 hypothetical prote
809 31 38.3 132 2 I40566 hypothetical prote
810 31 38.3 137 1 HSTE91 histone H2A.1 - Te
811 31 38.3 137 2 B72786 hypothetical prote
812 31 38.3 138 2 S41471 histone H2A.1 - Te
813 31 38.3 140 2 T36484 hypothetical prote
814 31 38.3 140 2 F81446 hypothetical prote
815 31 38.3 141 1 HTOR hemoglobin theta-1
816 31 38.3 141 2 H71975 ribosomal protein
817 31 38.3 141 2 D64530 ribosomal protein
818 31 38.3 141 2 S01772 Spec2d protein - s
819 31 38.3 142 1 HANE hemoglobin alpha c
820 31 38.3 142 1 HTHU hemoglobin theta-1
821 31 38.3 142 2 A27792 hemoglobin theta-1
822 31 38.3 146 2 B82700 hypothetical prote
823 31 38.3 159 2 B90361 hypothetical prote
824 31 38.3 160 2 S76328 hypothetical prote
825 31 38.3 160 2 A53293 flm3-region hypoth
826 31 38.3 163 2 T12561 glycine cleavage s
827 31 38.3 163 2 D39925 hypothetical prote
828 31 38.3 171 2 AH2855 conserved hypothet
829 31 38.3 171 2 E97632 hypothetical prote
830 31 38.3 173 2 F97105 HD family hydrolas
831 31 38.3 174 2 C82396 transcription regu
832 31 38.3 177 2 AB3221 oxidoreductase wit

833	31	38.3	178	2	PT0219	Ig kappa chain V-C	906	31	38.3	292	2	G95926	probable saccharid
834	31	38.3	180	2	S06616	chorion protein s1	907	31	38.3	293	2	AE3155	oxidoreductase Atu
835	31	38.3	181	2	H72853	fibroblast growth	908	31	38.3	295	2	C83244	conserved hypothet
836	31	38.3	183	2	H72572	hypothetical prote	909	31	38.3	295	2	A44064	hepatic leukemia f
837	31	38.3	183	2	AD2549	hypothetical prote	910	31	38.3	295	2	S58525	hepatic leukemia f
838	31	38.3	188	2	G95889	probable oxidoredu	911	31	38.3	295	2	H83961	L-serine dehydrata
839	31	38.3	188	2	AD0945	probable DNA-bindi	912	31	38.3	298	2	S72880	hypothetical prote
840	31	38.3	192	2	C71602	hypothetical prote	913	31	38.3	299	1	D70181	protein-export mem
841	31	38.3	196	2	I38022	hypothetical prote	914	31	38.3	300	2	F75586	urea/short-chain a
842	31	38.3	198	2	C32998	chorion protein S1	915	31	38.3	300	2	C70586	probable bex - Myc
843	31	38.3	199	2	C83761	hypothetical prote	916	31	38.3	309	2	C47057	homoserine kinase
844	31	38.3	203	2	T36403	probable tetR-fami	917	31	38.3	312	1	S46047	probable 3-methyl-
845	31	38.3	205	2	AH1907	transcription regu	918	31	38.3	313	2	D98132	hypothetical prote
846	31	38.3	206	2	F72670	hypothetical prote	919	31	38.3	324	2	AB0982	probable 2-hydroxy
847	31	38.3	208	2	G89783	hypothetical prote	920	31	38.3	326	2	S59101	NADH2 dehydrogenas
848	31	38.3	208	2	AH2535	hypothetical prote	921	31	38.3	328	2	B86030	probable dehydroge
849	31	38.3	209	2	S32681	hypothetical prote	922	31	38.3	328	2	C65154	probable 2-hydroxy
850	31	38.3	210	2	E97843	DNA-directed RNA p	923	31	38.3	328	2	F91183	probable dehydroge
851	31	38.3	211	2	A05123	DNA-(apurinic or a	924	31	38.3	331	1	A54932	zeta-crystallin /
852	31	38.3	211	2	AB3286	hypothetical prote	925	31	38.3	331	2	AG3140	hypothetical prote
853	31	38.3	211	2	T34501	monofunctional bio	926	31	38.3	331	2	E98147	hypothetical prote
854	31	38.3	212	2	F71634	hypothetical prote	927	31	38.3	331	2	B82341	NifR3/Smml family
855	31	38.3	212	2	AH1944	endonuclease III (928	31	38.3	333	2	T33311	hypothetical prote
856	31	38.3	214	2	I69804	phosphoglycerate m	929	31	38.3	334	2	T41713	hypothetical prote
857	31	38.3	214	2	B86639	acyl carrier prote	930	31	38.3	335	2	T25498	hypothetical prote
858	31	38.3	215	2	G82564	3-isopropylmalate	931	31	38.3	336	2	AB2525	hypothetical prote
859	31	38.3	218	2	E85021	hypothetical prote	932	31	38.3	339	2	E83496	hypothetical prote
860	31	38.3	220	2	A95956	hypothetical prote	933	31	38.3	340	1	RGFFBH	GTP-binding regula
861	31	38.3	221	2	B69861	hypothetical prote	934	31	38.3	341	2	C83067	hypothetical prote
862	31	38.3	221	2	G82427	alkaline phosphata	935	31	38.3	344	2	AE3228	agrocinopine synth
863	31	38.3	223	2	AH1815	hypothetical prote	936	31	38.3	344	2	S51948	hypothetical prote
864	31	38.3	224	2	T12797	two-component resp	937	31	38.3	345	1	B46113	protein kinase (EC
865	31	38.3	227	2	D90668	immunity protein d	938	31	38.3	349	1	TVCHMS	protein kinase (EC
866	31	38.3	229	2	F64754	probable xanthine	939	31	38.3	351	2	I39808	spore coat protein
867	31	38.3	229	2	G85518	probable oxidoredu	940	31	38.3	357	2	A72280	conserved hypothet
868	31	38.3	229	2	T40439	glucuronolactone r	941	31	38.3	359	2	E83251	probable aminotran
869	31	38.3	229	2	G83565	hypothetical prote	942	31	38.3	360	2	C64628	lipid A disacchari
870	31	38.3	231	2	S28703	probable bacteriop	943	31	38.3	360	2	D71888	lipid-a-disacchari
871	31	38.3	231	2	B69046	hypothetical prote	944	31	38.3	360	2	AE1931	hypothetical prote
872	31	38.3	232	2	T16258	precorrin-2 methyl	945	31	38.3	362	2	AE0019	3-dehydroquinat s
873	31	38.3	232	2	A87612	hypothetical prote	946	31	38.3	362	2	B87789	protein C34G6.5 [i
874	31	38.3	233	2	B64436	petR protein, prob	947	31	38.3	363	2	T22255	hypothetical prote
875	31	38.3	233	2	E86223	cobalamin biosynth	948	31	38.3	364	2	T23819	hypothetical prote
876	31	38.3	237	2	T25955	hypothetical prote	949	31	38.3	367	2	A24411	uroporphyrinogen d
877	31	38.3	238	2	S76860	hypothetical prote	950	31	38.3	367	2	G02786	uroporphyrinogen d
878	31	38.3	238	2	C71725	hypothetical prote	951	31	38.3	367	2	T10088	uroporphyrinogen d
879	31	38.3	245	2	F83270	amino-acid ABC tra	952	31	38.3	368	2	C29356	hydroxyproline-ric
880	31	38.3	245	2	B25852	probable nucleosid	953	31	38.3	369	2	T24022	hypothetical prote
881	31	38.3	247	1	JC1374	trypsin (EC 3.4.21	954	31	38.3	372	2	JC1069	nonstructure prote
882	31	38.3	247	2	B83816	carboxylesterase (955	31	38.3	372	2	C96631	hypothetical prote
883	31	38.3	247	2	F69868	glucose 1-dehydrog	956	31	38.3	372	2	T25629	hypothetical prote
884	31	38.3	248	2	H83111	hypothetical prote	957	31	38.3	374	2	C82142	alanine dehydrogen
885	31	38.3	251	2	A99183	tRNA pseudouridine	958	31	38.3	377	2	T00643	hypothetical prote
886	31	38.3	251	2	F83101	conserved hypothet	959	31	38.3	377	2	AB1552	zinc metalloprotei
887	31	38.3	252	2	S31147	endonuclease (EC 3	960	31	38.3	377	2	AD1194	N-acetylglucosamin
888	31	38.3	252	2	S74480	hypothetical prote	961	31	38.3	377	2	AD1194	N-acetylglucosamin
889	31	38.3	253	2	S56108	hypothetical prote	962	31	38.3	380	2	H81830	8-amino-7-oxononan
890	31	38.3	257	2	E95071	26S proteasome reg	963	31	38.3	380	2	C81194	8-amino-7-oxononan
891	31	38.3	259	2	G95890	tributylin esteras	964	31	38.3	384	2	G82976	probable rubredoxi
892	31	38.3	259	2	C97939	probable transcrip	965	31	38.3	384	2	S74806	hypothetical prote
893	31	38.3	259	2	AD0171	tributylin esteras	966	31	38.3	389	2	S67450	probable phosphate
894	31	38.3	261	2	D71838	probable methyltra	967	31	38.3	393	1	TVBE66	protein kinase (EC
895	31	38.3	265	2	I53413	nadh oxidoreductas	968	31	38.3	394	2	T31891	hypothetical prote
896	31	38.3	268	2	H82757	calcium sensor pro	969	31	38.3	396	2	S49592	cytochrome b-like
897	31	38.3	268	2	AC0497	hypothetical prote	970	31	38.3	396	2	S56954	protein YBR162c ho
898	31	38.3	269	2	H72484	probable haloacid	971	31	38.3	398	2	C84579	probable SET-domai
899	31	38.3	272	2	S00953	probable arylmalon	972	31	38.3	399	2	S34681	hypothetical prote
900	31	38.3	273	2	C87634	gag protein - frui	973	31	38.3	400	2	S30382	glycine hydroxymet
901	31	38.3	275	2	A75356	conserved hypothet	974	31	38.3	402	2	B95354	hypothetical prote
902	31	38.3	284	2	A43556	conserved hypothet	975	31	38.3	403	2	C75405	streptomycin biosy
903	31	38.3	285	1	T15252	homeotic protein H	976	31	38.3	406	2	G84774	hypothetical prote
904	31	38.3	285	2	A95392	hypothetical prote	977	31	38.3	412	2	F90004	serine hydroxymeth
905	31	38.3	291	2		probable LysR-fami	978	31	38.3	413	2	E84120	serine hydroxymeth

979 31 38.3 413 2 AC1392 glycine hydroxymet
980 31 38.3 413 2 AB1767 glycine hydroxymet
981 31 38.3 414 1 JQ1016 glycine hydroxymet
982 31 38.3 414 2 H81383 glycine hydroxymet
983 31 38.3 414 2 S18530 hypothetical prote
984 31 38.3 415 1 I40483 glycine hydroxymet
985 31 38.3 417 1 S00845 hepsin (EC 3.4.21.
986 31 38.3 417 2 G84963 glycine hydroxymet
987 31 38.3 417 2 H70174 glycine hydroxymet
988 31 38.3 419 2 T15199 hypothetical prote
989 31 38.3 421 2 T43534 transcription fact
990 31 38.3 421 2 T14773 hypothetical prote
991 31 38.3 423 1 DEBYMC malate dehydrogena
992 31 38.3 425 1 KRXL2A keratin, 64K type
993 31 38.3 427 2 F72341 glycine hydroxymet
994 31 38.3 427 2 S75210 glycine hydroxymet
995 31 38.3 427 2 AF2406 serine hydroxymeth
996 31 38.3 427 2 I51580 XFKH2 protein - Af
997 31 38.3 430 2 S12705 site-specific DNA-
998 31 38.3 430 2 JQ1050 deoxyfructosyl-glu
999 31 38.3 430 2 I39727 mannopine biosynth
1000 31 38.3 434 2 D72390 conserved hypothet

ALIGNMENTS

RESULT 1
A32378
probable ATP-dependent RNA helicase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change 09-Jul-2004
C;Accession: A32378
R;Leroy, P.; Alizari, P.; Sassoon, D.; Wolgemuth, D.; Fellous, M.
Cell 57, 549-559, 1989
A;Title: The protein encoded by a murine male germ cell-specific transcript is a putative
A;Reference number: A32378; MUID:89249320; PMID:2720782
A;Accession: A32378
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-660 <LER>
A;Cross-references: UNIPROT:P16381; UNIPARC:UPI0000003F94; GB:J04847; NID:g200388; PIDN:
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: Arp; DNA binding; nucleotide binding; P-loop
F;223-230/Region: nucleotide-binding motif A (P-loop)
F;342-347/Region: nucleotide-binding motif B
F;346-349/Region: DEAD motif

Query Match 59.3%; Score 48; DB 2; Length 660;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||:|||||
Db 455 SLEDFLYHEGYAC 467

RESULT 2
I84741
RNA helicase - mouse
N;Alternate names: RNA helicase ERH
C;Species: Mus musculus (house mouse)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: I84741; S56112
R;Gee, S.L.; Conboy, J.G.
Gene 140, 171-177, 1994
A;Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting h
A;Reference number: I49731; MUID:94192995; PMID:8144024
A;Accession: I84741
A;Status: translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-662 <RES>
A;Cross-references: UNIPROT:Q62167; UNIPARC:UPI0000044D80; GB:L25126; NID:g407995; PIDN:

A;Genetics: RES1
R;Sowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.
Biochem. J. 308, 839-846, 1995
A;Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of R
A;Reference number: S56112; MUID:97104282; PMID:8948440
A;Accession: S56112
A;Molecule type: mRNA
A;Residues: 1-229,'R',231-662 <SOW>
A;Cross-references: UNIPARC:UPI0000172EBF; GB:Z38117; NID:g1835121; PID:g1835122
A;Experimental source: strain c57bl/6; notochord
A;Genetics: SOW1
A;Note: the sequence is revised in GenBank entry MMDBRNAHL, release 117, (PID:1835122)
A;Note: the revised sequence is now identical to PIR accession I84741
C;Genetics: <RES1>
A;Gene: Eif4a-rs1; MGI:Ddx19
A;Cross-references: MGI:99526
C;Genetics: <SOW1>
A;Gene: MGI:Ddx3
A;Cross-references: MGI:103064
A;Map position: 1
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;224-231/Region: nucleotide-binding motif A (P-loop)
F;343-348/Region: nucleotide-binding motif B
F;347-350/Region: DEAD motif

Query Match 59.3%; Score 48; DB 1; Length 662;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||:|||||
Db 456 SLEDFLYHEGYAC 468

RESULT 3
A81710
ribosomal protein S1 TC0373 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: A81710
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: A81710
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <TET>
A;Cross-references: UNIPROT:P38016; UNIPARC:UPI00000578D3; GB:AE002304; GB:AE002160; NID:
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0373
C;Superfamily: Escherichia coli ribosomal protein S1

Query Match 55.6%; Score 45; DB 2; Length 536;
Best Local Similarity 63.6%; Pred. No. 8.8;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGY 12
|:|||||:
Db 512 SIKEFLAHGGH 522

RESULT 4
S13654
ATP-dependent RNA helicase - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S13654; S29676
R;Gururajan, R.; Perry-O'Keefe, H.; Melton, D.A.; Weeks, D.L.
Nature 349, 717-719, 1991

A;Title: The Xenopus localized messenger RNA An3 may encode an ATP-dependent RNA helicase
A;Reference number: S13654; MUID:91141586; PMID:1996140
A;Accession: S13654
A;Molecule type: mRNA
A;Residues: 1-697 <GUR>
A;Cross-references: UNIPROT:P24346; UNIPARC:UPI0000125AC2; EMBL:X57328; NID:g65059; PIDN:
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;260-272/Region: ATP binding #status predicted
F;265-272/Region: nucleotide-binding motif A (P-loop)
F;384-389/Region: nucleotide-binding motif B
F;386-398/Region: ATP binding #status predicted
F;388-391/Region: DEAD motif

Query Match 55.6%; Score 45; DB 1; Length 697;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLQEFLSHGGYVC 14
Db 497 ALEDFLYHEGYAC 509
:|:|:|:|:|:|

RESULT 5
JC5401
inositol 1,3,4-trisphosphate 5/6-kinase (EC 2.7.--.) [imported] - Arabidopsis thaliana
N;Alternate names: protein F5E19_100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Jul-1997 #sequence revision 18-Jul-1997 #text_change 09-Jul-2004
C;Accession: JC5401; T51508; T51954
R;Wilson, M.P.; Majerus, P.W.
Biochem. Biophys. Res. Commun. 232, 678-681, 1997

A;Title: Characterization of a cDNA encoding Arabidopsis thaliana inositol 1,3,4-trisphosphate
A;Reference number: JC5401; MUID:97271384; PMID:9126335
A;Accession: JC5401
A;Molecule type: mRNA

A;Residues: 1-319 <WIL>
A;Cross-references: UNIPROT:O81633; UNIPARC:UPI000009EB2A; GB:AF080173; NID:g3396078; PIDN:
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000

A;Reference number: Z25394
A;Accession: T51508
A;Molecule type: DNA
A;Residues: 1-196,'K',198-319 <SAT>
A;Cross-references: UNIPARC:UPI00000ACC41; EMBL:AL391147; NID:g9755718; PIDN:CAC01840.1;
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:

A;Map position: 5
A;Note: F5E19_100
C;Keywords: phosphotransferase

Query Match 54.3%; Score 44; DB 2; Length 319;
Best Local Similarity 63.6%; Pred. No. 7.7;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYV 13
Db 176 LQEFVNHGGVI 186
|:|:|:|:|:|

RESULT 6
A36220
transforming protein (s-myc) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Oct-1991 #sequence revision 13-Jan-1993 #text_change 09-Jul-2004
C;Accession: A36220

R;Sugiyama, A.; Kume, A.; Nemoto, K.; Lee, S.Y.; Asami, Y.; Nemoto, F.; Nishimura, S.; K
Proc. Natl. Acad. Sci. U.S.A. 86, 9144-9148, 1989

A;Title: Isolation and characterization of s-myc, a member of the rat myc gene family.
A;Reference number: A36220; MUID:90083224; PMID:2594755
A;Accession: A36220
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-429 <SUG>
A;Cross-references: UNIPROT:P23999; UNIPARC:UPI000012FAFE; GB:M29069; NID:g205553; PIDN:
C;Superfamily: myc transforming protein; myc transforming protein homology
C;Keywords: DNA binding; leucine zipper; nucleus; phosphoprotein
F;19-429/Domain: myc transforming protein homology <MYC>
F;398-426/Region: leucine zipper motif

Query Match 54.3%; Score 44; DB 2; Length 429;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYV 13
Db 227 SLEDFLSNSGYV 238
|:|:|:|:|:|

RESULT 7
T19954
hypothetical protein C45B11.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 05-Oct-2004
C;Accession: T19954

R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19202
A;Accession: T19954
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 1-293 <WIL>
A;Cross-references: UNIPROT:Q18639; UNIPARC:UPI0000077952; EMBL:Z74029; PIDN:CAA98431.1;
A;Experimental source: clone C45B11
C;Genetics:

A;Gene: CESP:C45B11.3
A;Map position: 5
A;Introns: 9/2; 66/3
C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 51.9%; Score 42; DB 2; Length 293;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGG 11
Db 208 SAMEFLSHGG 217
|:|:|:|:|:|

RESULT 8
T10544
inositol 1,3,4-trisphosphate 5/6-kinase homolog T12G13.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10544

R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10544
A;Molecule type: DNA

A;Residues: 1-338 <BEV>
A;Cross-references: UNIPROT:Q9SUG3; UNIPARC:UPI0000048910; EMBL:AL080252; GSPDB:GN00062,
A;Experimental source: cultivar Columbia; BAC clone T12G13
C;Genetics:

A;Gene: ATSP:T12G13.10
A;Map position: 4
A;Introns: 65/2; 92/3; 103/3; 144/1; 175/1; 215/1; 274/1; 292/3

Query Match 51.9%; Score 42; DB 2; Length 338;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LQEFLSHGG 11
Db 208 LQEFVNHGG 216
|:|:|:|:|:|

RESULT 9
B72313
glutamine synthetase - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: B72313
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: B72313
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <ARN>
A;Cross-references: UNIPROT:P36205; UNIPARC:UPI000012B707; GB:AE001758; GB:AE000512; NID
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0943
C;Superfamily: glutamate-ammonia ligase

Query Match 51.9%; Score 42; DB 2; Length 439;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 5 EFLSHGGY 12
Db 145 EFLDHGGY 152

RESULT 10
H90876
probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90876
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <HAY>
A;Cross-references: UNIPROT:Q8X2I0; UNIPARC:UPI00000D2AC7; GB:BA000007; PIDN:BA035407.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1984
C;Superfamily: phage lambda minor tail protein M

Query Match 51.2%; Score 41.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 7.3;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 LQEFLS-HGGY 12
Db 62 LEEFLSWHGGY 72

RESULT 11
T23345
hypothetical protein K05C4.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23345
R;Harris, B.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19729
A;Accession: T23345
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-436 <WIL>
A;Cross-references: UNIPROT:Q9XUU2; UNIPARC:UPI0000076837; EMBL:Z81564; PIDN:CAB04576.1;
A;Experimental source: clone K05C4
C;Genetics:
A;Gene: CESP:K05C4.11
A;Map position: 1
A;Introns: 34/1; 146/1; 174/1; 333/3; 369/1
C;Superfamily: Caenorhabditis elegans protein K05C4.11; LDL receptor ligand-binding repeat F;330-369/Domain: LDL receptor ligand-binding repeat homology <LDL>

Query Match 50.6%; Score 41; DB 2; Length 436;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 EFLSHGGYVC 14
Db 321 QFMCHGGQVC 330

RESULT 12
I38026
MLN 62 protein - human
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 07-Jul-2003
C;Accession: I38026; S60681
R;Tomasetto, C.; Regnier, C.H.; Moog-Lutz, C.; Mattei, M.G.; Chenard, M.P.; Lidereau, R. Genomics 28, 367-376, 1995
A;Title: Identification of four novel human genes amplified and overexpressed in breast cancer
A;Reference number: I37080; MUID:96039245; PMID:7490069
A;Accession: I38026
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-470 <RES>
A;Cross-references: UNIPARC:UPI0000073E24; EMBL:X80200; NID:g951276; PIDN:CAA56491.1; PIR
A;Note: submitted to the EMBL Data Library, July 1994
C;Genetics:
A;Gene: MLN62; CART1
C;Superfamily: TNF receptor-associated factor (TRAF); RING finger homology F;14-63/Domain: RING Finger homology <RRN>

Query Match 50.6%; Score 41; DB 2; Length 470;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQEFLSHGGYVC 14
Db 43 LQEFLSEGVEFKC 54

RESULT 13
G88925
protein T22H9.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G88925
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G88925
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-880 <STO>
A;Cross-references: UNIPROT:Q9TXN6; UNIPARC:UPI0000080264; GB:chr_V; PIDN:AAC69227.1; PIR
C;Genetics:
A;Gene: T22H9.2
A;Map position: 5

Query Match 50.0%; Score 40.5; DB 2; Length 880;
Best Local Similarity 40.0%; Pred. No. 87;

Matches 8; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

QY 2 SLOEFLSH-----GGYVC 14
:| ||| :| |||||

Db 135 NLDEFFTHIYEHQNGGYLC 154

RESULT 14
PQ0446
DNA-binding protein B2F - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PQ0446; JQ1679; S27764
R;Perisic, O.; Lam, E.
Plant Cell 4, 831-838, 1992
A;Title: A tobacco DNA binding protein that interacts with a light-responsive box II ele
A;Reference number: PQ0446; MUID:93005727; PMID:1392597
A;Accession: PQ0446
A;Molecule type: mRNA
A;Residues: 1-392 <PER>
A;Cross-references: UNIPROT:Q43326; UNIPARC:UPI000000ACF45; GB:S46942; NID:g258167; PIDN:
A;Experimental source: leaf, strain SR1
R;Gilmartin, P.M.; Memelink, J.; Hiratsuka, K.; Kay, S.A.; Chua, N.H.
Plant Cell 4, 839-849, 1992
A;Title: Characterization of a gene encoding a DNA binding protein with specificity for
A;Reference number: JQ1679; MUID:93005728; PMID:1392598
A;Accession: JQ1679
A;Molecule type: mRNA
A;Residues: 13-392 <GIL>
A;Cross-references: UNIPARC:UPI000016DE7F; GB:M93436; NID:g170270; PIDN:AAA34085.1; PID:
C;Keywords: DNA binding
F;75-103/Domain: alpha-helical rod <RD1>
F;115-124/Domain: alpha-helical rod <RD2>
F;138-150/Domain: alpha-helical rod <RD3>
F;159-174/Domain: alpha-helical rod <RD4>

Query Match 49.4%; Score 40; DB 2; Length 392;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEFLSHGGYVC 14
:||||| |:

Db 356 RDFLSHRGWTC 366

RESULT 15
A46394
suppressor protein SSL1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein L1531; protein YLR005w
C;Species: Saccharomyces cerevisiae
C;Date: 18-May-1994 #sequence_revision 19-Jul-1996 #text_change 31-Dec-2004
C;Accession: A46394; S64827
R;Yoon, H.; Miller, S.P.; Pabich, E.K.; Donahue, T.F.
Genes Dev. 6, 2463-2477, 1992
A;Title: SSL1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for transl
A;Reference number: A46394; MUID:94040711; PMID:1340463
A;Accession: A46394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <YOO>
A;Cross-references: UNIPROT:Q04673; UNIPARC:UPI0000053049; GB:Z17385; NID:g2695; PID:g26
R;Vandenbol, M.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64742
A;Accession: S64827
A;Molecule type: DNA
A;Residues: 1-461 <VAN>
A;Cross-references: UNIPARC:UPI0000053049; EMBL:Z73177; NID:g1360293; PID:g1360294; MIPS
A;Note: experimental_source strain S288C
C;Genetics:
A;Gene: SGD:SSL1
A;Cross-references: SGD:S0003995; MIPS:YLR005w
A;Map position: 12R

C;Superfamily: TFIIH basal transcription factor complex, subunit SSL1
C;Keywords: transmembrane protein
F;356-372/Domain: transmembrane #status predicted <TMM>

Query Match 49.4%; Score 40; DB 2; Length 461;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CSLOEFLSHGGYVC 14
|| | :||| |

Db 336 CSCHSKLVYGGYFC 349

RESULT 16
A83884
L-arabinose isomerase araA [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A83884
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83884
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-497 <STO>
A;Cross-references: UNIPROT:Q9KBQ2; UNIPARC:UPI0000125D50; GB:AP001513; GB:BA000004; NII
A;Experimental source: strain C-125
C;Genetics:
A;Gene: araA
C;Superfamily: L-arabinose isomerase

Query Match 49.4%; Score 40; DB 2; Length 497;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLOEFLSHGGY 12
:|:||| |||

Db 262 ALKEFLEEGY 272

RESULT 17
A71556
probable S1 ribosomal protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: A71556
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: A71556
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-569 <ARN>
A;Cross-references: UNIPROT:O84100; UNIPARC:UPI0000047BDA; GB:AE001284; GB:AE001273; NII
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: rsl
C;Superfamily: Escherichia coli ribosomal protein S1

Query Match 49.4%; Score 40; DB 2; Length 569;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SLOEFLSHGG 11
|:|||||

Db 546 SIKEFLVHGG 555

RESULT 18
T27429

hypothetical protein Y79H2A.11 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T27429
R;Matthews, L.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z20365
A/Accession: T27429
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-643 <WIL>
A/Cross-references: UNIPROT:Q95QC4; UNIPARC:UPI000017BCBE; EMBL:AL110501; NID:e1542357;
A/Experimental source: clone Y79H2A
C/Genetics:
A/Gene: CESP:Y79H2A.11
A/Introns: 63/1; 365/2; 443/2; 502/2; 562/3; 611/3

Query Match 49.4%; Score 40; DB 2; Length 643;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
| : | | | | |
Db 274 SVDQFEDGGGYVC 286

RESULT 19
D56695
transducin-like enhancer-of-split homolog TLE-3 - human
C/Species: Homo sapiens (man)
C/Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 05-Oct-2004
C/Accession: D56695
R;Stifani, S.; Blaumueeller, C.M.; Redhead, N.J.; Hill, R.E.; Artavanis-Tsakonas, S.
Nature Genet. 2, 119-127, 1992
A/Title: Human homologs of a Drosophila enhancer of split gene product define a novel fa
A/Reference number: A56695; MUID:93265135; PMID:1303260
A/Accession: D56695
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-772 <STI>
A/Cross-references: UNIPROT:Q04726; UNIPARC:UPI0000163B5B; GB:M99438; NID:g307513; PIDN:
C/Genetics:
A/Gene: GDB:TLE3; ESG; ESG3
A/Cross-references: GDB:228049; OMIM:600190
A/Map position: 15pter-15qter
C/Keywords: nucleus
F/482-514/Domain: WD repeat homology <WD1>
F/528-561/Domain: WD repeat homology <WD2>
F/614-647/Domain: WD repeat homology <WD3>
F/696-729/Domain: WD repeat homology <WD4>
F/734-770/Domain: WD repeat homology <WD5>

Query Match 49.4%; Score 40; DB 2; Length 772;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 LSHGGYVC 14
| | | | | |
Db 483 LSHGGYVC 490

RESULT 20
D90854
minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: D90854
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D90854

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <HAY>
A/Cross-references: UNIPROT:Q8X5F5; UNIPARC:UPI00000D2966; GB:BA000007; PIDN:BAB35227.1;
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs1804
C/Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3 LQEFLS-HGGY 12
| | | | | | | |
Db 62 LDEFLSWHGGY 72

RESULT 21
G90908
probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: G90908
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: G90908
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <HAY>
A/Cross-references: UNIPROT:Q8X5F5; UNIPARC:UPI00000D2966; GB:BA000007; PIDN:BAB35662.1;
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs2239
C/Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3 LQEFLS-HGGY 12
| | | | | | | |
Db 62 LDEFLSWHGGY 72

RESULT 22
B90823
probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: B90823
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: B90823
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <HAY>
A/Cross-references: UNIPROT:Q8X3G6; UNIPARC:UPI00000D298C; GB:BA000007; PIDN:BAB34977.1;
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs1554
C/Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 283 CSRQHIFEHNQYMC 296

RESULT 27

C90154

primase SS00140 [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C;Accession: C90154

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: C90154

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-532 <KUR>

A;Cross-references: UNIPROT:Q980Y2; UNIPARC:UPI0000064196; GB:AE006641; NID:gl13813270; E

C;Genetics:

A;Gene: SS00140

Query Match 48.1%; Score 39; DB 2; Length 532;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYV 13
|:|: :||| ||
Db 93 CTLEWKISHGRYV 105

RESULT 28

T35002

probable respiratory chain oxidoreductase - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35002

R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1998

A;Reference number: Z21564

A;Accession: T35002

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-646 <SEE>

A;Cross-references: UNIPROT:Q9ZBV8; UNIPARC:UPI00000DAE34; EMBL:AL034443; PIDN:CRAA22364.

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SGOEPB:SC4B5.10c

C;Superfamily: hydrogen dehydrogenase alpha chain; NADH dehydrogenase (ubiquinone) I cha

Query Match 48.1%; Score 39; DB 2; Length 646;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGY 12
||: :||| ||
Db 257 SLDDYRAHGGY 267

RESULT 29

B83725

hypothetical protein BH0602 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: B83725

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: B83725

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-652 <STO>

A;Cross-references: UNIPROT:Q9KF83; UNIPARC:UPI00000C3921; GB:AP001509; GB:BA000004; NID

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0602

Query Match 48.1%; Score 39; DB 2; Length 652;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LQEFLSHGGY 12
|:|: :||| ||
Db 168 LPGFLNHGGY 177

RESULT 30

AG0265

alcohol dehydrogenase (EC 1.1.1.1) [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 31-Dec-2004

C;Accession: AG0265

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; E
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E
Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AG0265

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-891 <KUR>

A;Cross-references: UNIPROT:Q8ZEJ0; UNIPARC:UPI00000CD8B1; GB:AL590842; PIDN:CAC90987.1;

C;Genetics:

A;Gene: adhE

C;Superfamily: bifunctional aldehyde/alcohol dehydrogenase; aldehyde dehydrogenase homol

C;Keywords: oxidoreductase

Query Match 48.1%; Score 39; DB 2; Length 891;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEFLSHGGYV 13
:|:|:|:|:|:
Db 264 ERFASHGGYL 273

RESULT 31

AE0650

alcohol dehydrogenase (EC 1.1.1.1) [imported] - Salmonella enterica subsp. enterica sero

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 31-Dec-2004

C;Accession: AE0650

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AE0650

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-892 <PAR>

A;Cross-references: UNIPARC:UPI0000059F5E; GB:AL513382; PIDN:CAD08384.1; PID:gl6502428;

C;Genetics:

A;Gene: adh

C;Superfamily: bifunctional aldehyde/alcohol dehydrogenase; aldehyde dehydrogenase homol

C;Keywords: oxidoreductase

Query Match 48.1%; Score 39; DB 2; Length 892;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 QEFLSHGGYV 13
: |||||:
Db 264 ERFASHGGYM 273

RESULT 32
S19421
ATP-dependent permease ADP1 precursor - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YCR011c; protein YCR105
C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: S19421; S40914
R;Goffeau, A.; Purnelle, B.; Skala, J.
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19420
A;Accession: S19421
A;Molecule type: DNA
A;Residues: 1-1049 <GOF>
A;Cross-references: UNIPROT:P25371; UNIPARC:UPI00001255FD; EMBL:X59720; NID:g1907116; PI
R;Purnelle, B.; Skala, J.; Goffeau, A.
Yeast 7, 867-872, 1991
A;Title: The product of the YCR105 gene located on the chromosome III from Saccharomyces
A;Reference number: S40914; MUID:92160395; PMID:1789009
A;Accession: S40914
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1049 <PUR>
A;Cross-references: UNIPARC:UPI00001255FD
R;Skala, J.; Purnelle, B.; Goffeau, A.
Yeast 8, 409-417, 1992
A;Title: The complete sequence of a 10.8 kb segment distal of SUF2 on the right arm of c
K genes.
A;Reference number: S25353; MUID:92327849; PMID:1626432
A;Contents: annotation
C;Genetics:
A;Gene: SGD:ADP1; MIPS:YCR011c
A;Cross-references: SGD:S0000604; MIPS:YCR011c
A;Map position: 3R
C;Superfamily: ATP-dependent permease ADP1; ATP-binding cassette homology
C;Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1049/Product: ATP-dependent permease ADP1 #status predicted <MAT>
F;26-324/Domain: extracellular #status predicted <EXT>
F;325-341/Domain: transmembrane #status predicted <TM1>
F;406-607/Domain: ATP-binding cassette homology <ABC>
F;423-430/Domain: nucleotide-binding motif A (P-loop)
F;550-557/Region: nucleotide-binding motif B
F;794-810/Domain: transmembrane #status predicted <TM2>
F;829-845/Domain: transmembrane #status predicted <TM3>
F;878-894/Domain: transmembrane #status predicted <TM4>
F;909-925/Domain: transmembrane #status predicted <TM5>
F;938-954/Domain: transmembrane #status predicted <TM6>
F;1025-1041/Domain: transmembrane #status predicted <TM7>
F;50,114,165,221/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;429/Binding site: ATP (Lys) #status predicted

Query Match 48.1%; Score 39; DB 1; Length 1049;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LQEFSLSHGGYVC 14
: ||||: |||:
Db 618 VSEFLRNEGYYIC 629

RESULT 33
BVBYA1
3-dehydroquinase synthase (EC 4.2.3.4) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YD9302.02; protein YDR127w
N;Contains: 3-dehydroquinase synthase (EC 4.2.1.10); 3-dehydroquinase synthase (EC 4-
ikimate kinase (EC 2.7.1.71)
C;Species: Saccharomyces cerevisiae

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A32519; S51854; S42150; S42151
R;Duncan, K.; Edwards, R.M.; Coggins, J.R.
Biochem. J. 246, 375-386, 1987
A;Title: The pentafunctional AROM enzyme of Saccharomyces cerevisiae is a mosaic of mono
A;Reference number: A32519; MUID:88076802; PMID:2825635
A;Accession: A32519
A;Molecule type: DNA
A;Residues: 1-1588 <DUN>
A;Cross-references: UNIPROT:P08566; UNIPARC:UPI0000052F13; EMBL:X06077; NID:g3380; PIDN:
R;Oliver, K.; Harris, D.
submitted to the EMBL Data Library, February 1995
A;Reference number: S51853
A;Accession: S51854
A;Molecule type: DNA
A;Residues: 1-1588 <OLI>
A;Cross-references: UNIPARC:UPI0000052F13; EMBL:Z48179; NID:g665657; PIDN:CAA88208.1; PI
R;Duncan, K.; Edwards, R.M.; Coggins, J.R.
FEBS Lett. 241, 83-88, 1988
A;Title: The Saccharomyces cerevisiae ARO1 gene. An example of the co-ordinate regulatio
A;Reference number: S42150; MUID:89065111; PMID:2848727
A;Accession: S42150
A;Molecule type: DNA
A;Residues: 1-44;1557-1588 <DUW>
A;Cross-references: UNIPARC:UPI0000168B07; UNIPARC:UPI0000168B08; EMBL:X13802; EMBL:X138
C;Genetics:
A;Gene: SGD:ARO1; MIPS:YDR127w
A;Cross-references: SGD:S0002534; MIPS:YDR127w
A;Map position: 4R
C;Superfamily: pentafunctional Arom protein; 3-dehydroquinase dehydratase homology; 3-dc
homology; shikimate kinase homology
C;Keywords: aromatic amino acid biosynthesis; ATP; carbon-oxygen lyase; hydro-lyase; mul
F;7-380/Domain: 3-dehydroquinase synthase homology <DQS>
F;417-863/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>
F;887-1008/Domain: shikimate kinase homology <SKI>
F;1078-1291/Domain: 3-dehydroquinase dehydratase homology <DQD>
F;1362-1572/Domain: shikimate dehydrogenase homology <SKD>

Query Match 48.1%; Score 39; DB 1; Length 1588;
Best Local Similarity 58.3%; Pred. No. 2.9e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHGGYV 13
: |||: |||||
Db 981 ALKDFASSGGYV 992

RESULT 34
TI9130
hypothetical protein C09F9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19130
R;Smye, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19078
A;Accession: T19130
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1642 <WIL>
A;Cross-references: UNIPROT:O62055; UNIPARC:UPI000007E384; EMBL:Z81465; PIDN:CAB03861.1
A;Experimental source: clone C09F9
C;Genetics:
A;Gene: CESP:C09F9.2
A;Map position: 2
A;Introns: 75/3; 103/2; 220/1; 552/1; 1500/3; 1517/1; 1578/3

Query Match 48.1%; Score 39; DB 2; Length 1642;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLQEFSLSHGGYVC 14
||| | | |:
||| | | |:

Db 494 CSLNETLNECGRVC 507

RESULT 35

E86189

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: E86189

R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E86189

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1878 <STO>

A;Cross-references: UNIPROT:Q9SYJ7; UNIPARC:UPI000009DC9E; GB:AE005172; NID:g4836907; PI

C;Genetics:

A;Map position: 1

Query Match 47.5%; Score 38.5; DB 2; Length 1878;

Best Local Similarity 69.2%; Pred. No. 4.1e+02;

Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3 LQEFL-SHGGYVC 14

Db 1300 LQEFLEKHGGVRC 1312

RESULT 36

T03751

hypothetical protein A - slime mold (Physarum polycephalum) mitochondrion integrated pla

C;Species: mitochondrion Physarum polycephalum

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T03751

R;Nakagawa, C.C.; Jones, E.P.; Miller, D.L.

Curr. Genet. 33, 178-187, 1998

A;Title: Mitochondrial DNA rearrangements associated with mF plasmid integration and pla

A;Reference number: Z15055; MUID:98177147; PMID:9508792

A;Accession: T03751

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-94 <NAK>

A;Cross-references: UNIPROT:Q20627; UNIPARC:UPI000008B89C; EMBL:AF012249; NID:g2435419;

A;Experimental source: strain aux2-S

C;Genetics:

A;Genome: mitochondrion integrated plasmid mF

C;Keywords: mitochondrion

Query Match 46.9%; Score 38; DB 2; Length 94;

Best Local Similarity 42.9%; Pred. No. 24;

Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14

Db 77 CYIDKFCSIMGYIC 90

RESULT 37

A82985

hypothetical protein PA5285 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: A82985

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bn

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: A82985

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-105 <STO>

A;Cross-references: UNIPROT:Q9HTR9; UNIPARC:UPI00000C5F89; GB:AE004941; GB:AE004091; NID

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA5285

Query Match 46.9%; Score 38; DB 2; Length 105;

Best Local Similarity 58.3%; Pred. No. 27;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYV 13

Db 70 AMEEFLSRGGKV 81

RESULT 38

S46445

YMF46 protein - Acanthamoeba castellanii mitochondrion

C;Species: mitochondrion Acanthamoeba castellanii

C;Date: 26-Dec-1994 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004

C;Accession: S46445; S53825

R;Loneragan, K.M.; Gray, M.W.

J. Mol. Biol. 239, 476-499, 1994

A;Title: The ribosomal RNA gene region in Acanthamoeba castellanii mitochondrial DNA. A

A;Reference number: S46445; MUID:94275838; PMID:8006963

A;Accession: S46445

A;Molecule type: DNA

A;Residues: 1-142 <LON>

A;Cross-references: UNIPROT:Q37060; UNIPARC:UPI000008BB47; EMBL:U03732; NID:g495327; PID

R;Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.

J. Mol. Biol. 245, 522-537, 1995

A;Title: The mitochondrial DNA of the amoeboid protozoon, Acanthamoeba castellanii: comp

A;Reference number: S53825; MUID:95147275; PMID:7844823

A;Accession: S53825

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-142 <BUR>

A;Cross-references: UNIPARC:UPI000008BB47; GB:U12386; NID:g562028; PIDN:AAD11817.1; PID:

A;Experimental source: strain Neff; ATCC 30010

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC6

C;Keywords: mitochondrion

Query Match 46.9%; Score 38; DB 2; Length 142;

Best Local Similarity 38.5%; Pred. No. 37;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14

Db 52 AIKOFFGHGSVVC 64

RESULT 39

JE0136

lectin precursor - common snowdrop

C;Species: Galanthus nivalis (common snowdrop)

C;Date: 02-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 31-Dec-2004

C;Accession: JE0136; JC5675

R;Zhou, Y.; Tian, Y.; Tao, H.; Mang, K.

Chinese J. Biotechnol. 12, 495-498, 1996

A;Title: Molecular cloning and sequence analysis of snowdrop (Galanthus nivalis) lectin

A;Reference number: JE0136

A;Accession: JE0136

A;Molecule type: DNA
A;Residues: 1-160 <ZHO>
A;Cross-references: UNIPARC:UPI0000179130
A;Accession: JC5675
A;Molecule type: DNA
A;Residues: 1-160 <ZH2>
A;Cross-references: UNIPARC:UPI0000179130
A;Experimental source: leaf
C;Superfamily: lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;27-131/Product: lectin #status predicted <MAT>
F;132-160/Domain: carboxyl-terminal cleaved off #status predicted <CTC>

Query Match 46.9%; Score 38; DB 2; Length 160;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYV 13
|:|:|:|:|:
Db 37 STGEFLNYGGFV 48

RESULT 40
E64690
response regulator - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: E64690
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64690
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-213 <TOM>
A;Cross-references: UNIPROT:O25918; UNIPARC:UPI00000D3157; GB:AE000636; GB:AE000511; NID
C;Superfamily: ompR protein; response regulator homology
C;Keywords: phosphoprotein
F;5-115/Domain: response regulator homology <RRH>
F;53/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 46.9%; Score 38; DB 2; Length 213;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGY 12
|:|:|:|:|:
Db 17 SVKEFLEHLGY 27

RESULT 41
E82780
endonuclease III XF0647 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82780
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82780
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <SIM>
A;Cross-references: UNIPROT:Q9PFL0; UNIPARC:UPI00000C24A9; GB:AE003909; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0647
C;Superfamily: apurinic/apyrimidinic endonuclease III
C;Keywords: 4Fe-4S; metalloprotein
F;193,200,203,209/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 46.9%; Score 38; DB 2; Length 218;
Best Local Similarity 47.4%; Pred. No. 57;
Matches 9; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 4 QEFLS-----HGGYVC 14
|:|:|:|:|:
Db 175 QEFLKDAHHWLILHGRYVC 193

RESULT 42
D75275
endonuclease III - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: D75275
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; , M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75275
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-259 <WHI>
A;Cross-references: UNIPROT:Q9RRQ0; UNIPARC:UPI00000D3FB7; GB:AE002073; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2438
A;Map position: 1

Query Match 46.9%; Score 38; DB 2; Length 259;
Best Local Similarity 77.8%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 FLSHGGYVC 14
|:|:|:|:|:
Db 209 FLSHGRQVC 217

RESULT 43
AH2340
transcription regulator all4279 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2340
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2340
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-295 <KUR>
A;Cross-references: UNIPROT:Q8YPB9; UNIPARC:UPI00000CEA96; GB:BA000019; PIDN:BAB75978.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4279

Query Match 46.9%; Score 38; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLQEFLSHG 10
:|||||:
Db 29 AMQEFLLTHG 37

RESULT 44
T01266
starch synthase DULL1 - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01266
R;Gao, M.; Wanat, J.; Stinard, P.S.; James, M.G.; Myers, A.M.
Plant Cell 10, 399-412, 1998
A;Title: Characterization of dull1, a maize gene coding for a novel starch synthase.
A;Reference number: Z14287; MUID:98169346; PMID:9501113
A;Accession: T01266
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-310 <GAO>
A;Cross-references: UNIPROT:O64924; UNIPARC:UPI00000A655C; EMBL:AF023160; NID:G3057121;
C;Genetics:
A;Gene: dull1
A;Map position: 10
A;Note: Intron positions not resolved (incomplete sequence)

Query Match 46.9%; Score 38; DB 2; Length 310;
Best Local Similarity 61.5%; Pred. No. 81;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||| |||||
Db 294 SLSRMLSSGMYVC 306

RESULT 45
D83708
phosphonates transport system (phosphate-binding protein) BH0468 [imported] - Bacillus h
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: D83708
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <STO>
A;Cross-references: UNIPROT:Q9KFL1; UNIPARC:UPI00000D73BA; GB:AP001508; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0468

Query Match 46.9%; Score 38; DB 2; Length 323;
Best Local Similarity 63.6%; Pred. No. 85;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGY 12
|:|||||
Db 282 SMQEFLESNGY 292

RESULT 46

E71363
hypothetical protein TP0118 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: E71363
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
r, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: E71363
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-421 <COL>
A;Cross-references: UNIPROT:O83155; UNIPARC:UPI00001392CC; GB:AE001197; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0118

Query Match 46.9%; Score 38; DB 2; Length 421;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
||| | : || |
Db 26 CSLGEHIVRGAAC 39

RESULT 47
E64657
probable glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) - Helicobacter pylori (strain
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: E64657
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.M.; Venter, ,
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64657
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-425 <TOM>
A;Cross-references: UNIPROT:P56110; UNIPARC:UPI000012AF23; GB:AE000616; GB:AE000511; NID
C;Superfamily: glucose-6-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 46.9%; Score 38; DB 2; Length 425;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 QEFLSHGGYVC 14
:|||||:
Db 69 REFLAHISYLC 79

RESULT 48
A33504
glutamate dehydrogenase (NADp) (EC 1.4.1.4) - Salmonella typhimurium
N;Alternate names: glutamic dehydrogenase; NADP-specific glutamate dehydrogenase
C;Species: Salmonella typhimurium
C;Date: 08-Dec-1989 #sequence_revision 23-Mar-1995 #text_change 05-Oct-2004
C;Accession: A33504
R;Bansal, A.; Dayton, M.A.; Zalkin, H.; Colman, R.F.
J. Biol. Chem. 264, 9827-9835, 1989
A;Title: Affinity labeling of a glutamyl peptide in the coenzyme binding site of NADP(+)
enosine 2',5'-bisphosphate.
A;Reference number: A33504; MUID:89255551; PMID:2656714
A;Accession: A33504
A;Molecule type: DNA

A;Residues: 1-447 <BAN>												
A;Cross-references: UNIPROT:PI5111; UNIPARC:UPI0000129306; GB:M24021; GB:J04814												
C;Genetics:												
A;Gene: gdh												
A;Map position: 27 min												
C;Superfamily: glutamate dehydrogenase												
C;Keywords: homohexamer; NADP; oxidoreductase												
F;128/Binding site: substrate (Lys) #status predicted												
Query Match 46.9%; Score 38; DB 1; Length 447;												
Best Local Similarity 66.7%; Pred. No. 1.2e+02;												
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;												
Qy 1 CSLQEFLSH 9												
: :												
Db 5 CSLESFLNH 13												
RESULT 49												
AF0710												
NADP-specific glutamate dehydrogenase [imported] - Salmonella enterica subsp. enterica s												
C;Species: Salmonella enterica subsp. enterica serovar Typhi												
A;Note: this species has also been called Salmonella typhi												
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004												
C;Accession: AF0710												
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,												
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,												
, S.; Moule, S.; O'Gaora, P.												
Nature 413, 848-852, 2001												
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;												
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov												
A;Reference number: AB0502; MUID:21534947; PMID:11677608												
A;Accession: AF0710												
A;Status: preliminary												
A;Molecule type: DNA												
A;Residues: 1-447 <PAR>												
A;Cross-references: UNIPARC:UPI0000059DA7; GB:AL513382; PIDN:CAD02055.1; PID:gl6502892;												
C;Genetics:												
A;Gene: STY1815												
C;Superfamily: glutamate dehydrogenase												
Query Match 46.9%; Score 38; DB 2; Length 447;												
Best Local Similarity 66.7%; Pred. No. 1.2e+02;												
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;												
Qy 1 CSLQEFLSH 9												
: :												
Db 5 CSLESFLNH 13												
RESULT 50												
E83130												
conserved hypothetical protein PA4132 [imported] - Pseudomonas aeruginosa (strain PA01)												
C;Species: Pseudomonas aeruginosa												
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004												
C;Accession: E83130												
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B												
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,												
.; Lory, S.; Olson, M.V.												
Nature 406, 959-964, 2000												
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho												
A;Reference number: A82950; MUID:20437337; PMID:10984043												
A;Accession: E83130												
A;Status: preliminary												
A;Molecule type: DNA												
A;Residues: 1-471 <STO>												
A;Cross-references: UNIPROT:Q9HWP9; UNIPARC:UPI00000C5C29; GB:AE004829; GB:AE004091; NID												
A;Experimental source: strain PA01												
C;Genetics:												
A;Gene: PA4132												
Query Match 46.9%; Score 38; DB 2; Length 471;												
Best Local Similarity 45.5%; Pred. No. 1.2e+02;												

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-53
Perfect score: 81
Sequence: 1 CSLQEFLSHGGYVC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	60.5	488	Q6CNV5_KLULA	Q6cnv5 kluyveromyc
2	48	59.3	362	Q5S4N1_HUMAN	Q5s4n1 homo sapien
3	48	59.3	365	Q9FFN6_ARATH	Q9ffn6 arabidopsis
4	48	59.3	367	Q8GUG6_ARATH	Q8gug6 arabidopsis
5	48	59.3	480	Q4R9A4_MACFA	Q4r9a4 macaca fasc
6	48	59.3	658	1 DDX3Y_PONPY	Q5rf43 pongo pygma
7	48	59.3	660	1 DDX3Y_HUMAN	Q15523 homo sapien
8	48	59.3	660	1 DDX3Y_PANTR	Q6gvm6 pan troglod
9	48	59.3	660	1 PL10_MOUSE	P16381 mus musculu
10	48	59.3	661	1 DDX3X_HUMAN	O00571 homo sapien
11	48	59.3	661	1 DDX3X_MOUSE	Q62167 mus musculu
12	48	59.3	662	2 Q5JSI3_HUMAN	Q5jbi3 homo sapien
13	48	59.3	662	2 Q8K5D5_MESAU	Q8k5d5 mesocricetu
14	48	59.3	674	2 Q59GX6_HUMAN	Q59gx6 homo sapien
15	46	56.8	202	2 Q8QH15_TYRTY	Q8qh15 tyrannus ty
16	46	56.8	358	2 Q7XK97_ORYSA	Q7xk97 oryza sativ
17	46	56.8	666	2 Q69BL0_MANSE	Q69bl0 manduca sex
18	45	55.6	363	2 Q93ZD2_ARATH	Q93zd2 arabidopsis
19	45	55.6	363	2 Q9C5A2_ARATH	Q9c5a2 arabidopsis
20	45	55.6	376	2 Q5ER66_MONAL	Q5er66 monopterus
21	45	55.6	448	2 Q518T7_ENTHI	Q518t7 entamoeba h
22	45	55.6	570	1 RS1_CHLMU	P38016 chlamydia m
23	45	55.6	611	2 Q4KMN1_MOUSE	Q4kmn1 mus musculu
24	45	55.6	651	2 Q5F491_CHICK	Q5f491 gallus gall
25	45	55.6	672	2 Q8BIR2_MOUSE	Q8bir2 mus musculu
26	45	55.6	695	2 Q52L23_XENLA	Q52l23 xenopus lae
27	45	55.6	697	1 AN3_XENLA	P24346 xenopus lae
28	45	55.6	697	2 Q7ZXJ0_XENLA	Q7zxj0 xenopus lae
29	45	55.6	699	2 Q6P4J3_XENTR	Q6p4j3 xenopus tro
30	45	55.6	948	2 Q51FR1_ENTHI	Q51fr1 entamoeba h
31	44.5	54.9	651	2 Q9FJE2_ARATH	Q9fje2 arabidopsis

32	44	54.3	266	2	Q5DELO_SCHJA	Q5del0 schistosoma
33	44	54.3	319	2	O81633_ARATH	O81633 arabidopsis
34	44	54.3	319	2	Q9SBA5_ARATH	Q9sba5 arabidopsis
35	44	54.3	342	2	Q7XBW0_ORYSA	Q7xbw0 oryza sativ
36	44	54.3	342	2	Q84Y01_MAIZE	Q84y01 zea mays (m
37	44	54.3	342	2	Q8W3H7_ORYSA	Q8w3h7 oryza sativ
38	44	54.3	348	2	Q92M44_RHIME	Q92m44 rhizobium m
39	44	54.3	388	2	Q98E45_RHILO	Q98e45 rhizobium l
40	44	54.3	393	2	Q95K70_MACFA	Q95k70 macaca fasc
41	44	54.3	413	2	Q6AJD7_DESPS	Q6ajd7 desulfotale
42	44	54.3	429	1	MYCS_RAT	P23999 rattus norv
43	44	54.3	430	2	Q7N4F2_PHOLL	Q7n4f2 photorhabdu
44	44	54.3	489	2	Q5R8M2_PONPY	Q5r8m2 pongo pygma
45	44	54.3	490	2	Q6UX62_HUMAN	Q6ux62 homo sapien
46	44	54.3	490	2	Q5VX71_HUMAN	Q5vx71 homo sapien
47	44	54.3	701	2	Q4L8P7_STAHI	Q4l8p7 staphylococ
48	43	53.1	123	2	Q07363_CHLTR	Q07363 chlamydia t
49	43	53.1	219	2	Q4RU12_TETNG	Q4ru12 tetraodon n
50	43	53.1	469	2	Q759D9_ASHGO	Q759d9 ashbya goss
51	43	53.1	490	2	Q8BH32_MOUSE	Q8bh32 m mus muscu
52	43	53.1	492	2	Q8VC43_MOUSE	Q8vc43 mus musculu
53	43	53.1	516	2	Q608U7_METCA	Q608u7 methylococc
54	43	53.1	521	2	Q7PYM7_ANOGA	Q7pym7 anopheles g
55	43	53.1	666	2	Q5LXF8_SILPO	Q5lxf8 silicibacte
56	42	51.9	93	2	Q7NJV6_GLOVI	Q7njy6 gloeobacter
57	42	51.9	292	2	Q60ML3_CAEBR	Q60ml3 caenorhabdi
58	42	51.9	293	2	Q18639_CAEEL	Q18639 caenorhabdi
59	42	51.9	308	2	Q88JZ1_PSEPK	Q88jz1 pseudomonas
60	42	51.9	333	2	Q7XHG0_ORYSA	Q7xhg0 oryza sativ
61	42	51.9	333	2	Q9AYJ2_ORYSA	Q9ayj2 oryza sativ
62	42	51.9	338	2	Q9SUG3_ARATH	Q9sug3 arabidopsis
63	42	51.9	353	2	Q93VQ8_ARATH	Q93vq8 arabidopsis
64	42	51.9	355	2	Q6K7B8_ORYSA	Q6k7b8 oryza sativ
65	42	51.9	357	2	Q75GI4_ORYSA	Q75gi4 oryza sativ
66	42	51.9	357	2	Q8GSI5_ORYSA	Q8gsi5 oryza sativ
67	42	51.9	439	1	GLNA_THEMA	P36205 thermotoga
68	42	51.9	615	2	Q7V9F0_PROMA	Q7v9f0 prochloroco
69	42	51.9	626	2	Q7U3A8_SYNXP	Q7u3a8 synchococc
70	42	51.9	629	2	Q7V3R9_PROMM	Q7v3r9 prochloroco
71	42	51.9	789	2	Q8TXR3_METKA	Q8txr3 methanopyru
72	42	51.9	1871	2	Q67FT1_MESMA	Q67ft1 mesobuthus
73	41.5	51.2	113	2	Q8X2I0_ECO57	Q8x2i0 escherichia
74	41.5	51.2	113	2	Q6H9T6_9CAUD	Q6h9t6 phage phi 4
75	41.5	51.2	2043	2	Q54C50_DICDI	Q54c50 dictyosteli
76	41	50.6	100	2	Q9P1H9_HUMAN	Q9plh9 homo sapien
77	41	50.6	143	2	Q56C40_9CAUD	Q56c40 enterobacte
78	41	50.6	194	2	Q91G7_9TYRA	Q91g7 elaeinia fla
79	41	50.6	216	2	Q804U8_SAYPH	Q804u8 sayornis ph
80	41	50.6	274	2	Q5XIK0_RAT	Q5xik0 rattus norv
81	41	50.6	280	2	Q8STN6_ENCCU	Q8stn6 encephalito
82	41	50.6	320	2	Q88HI1_PSEPK	Q88hi1 pseudomonas
83	41	50.6	352	2	Q7P0Z9_CHRVO	Q7p0z9 chromobacte
84	41	50.6	354	2	Q8BJX2_MOUSE	Q8bjx2 mus musculu
85	41	50.6	363	2	Q4Q117_LEIMA	Q4ql17 leishmania
86	41	50.6	423	2	Q5CUE9_CRYPV	Q5cue9 cryptospori
87	41	50.6	436	2	Q9XUU2_CAEEL	Q9xuu2 caenorhabdi
88	41	50.6	470	1	TRAF4_HUMAN	Q9buz4 homo sapien
89	41	50.6	470	1	TRAF4_MOUSE	Q61382 mus musculu
90	41	50.6	470	2	Q8BHD9_MOUSE	Q8bhd9 m mus muscu
91	41	50.6	470	2	Q6DFG9_XENLA	Q6dfg9 xenopus lae
92	41	50.6	470	2	Q6DEV2_XENTR	Q6dev2 xenopus tro
93	41	50.6	470	2	Q68F27_XENLA	Q68f27 xenopus lae
94	41	50.6	470	2	Q7TLK2_BRARE	Q7tlk2 brachydanio
95	41	50.6	471	2	Q4RRL8_TETNG	Q4rrl8 tetraodon n
96	41	50.6	503	2	Q75EB7_ASHGO	Q75eb7 ashbya goss
97	41	50.6	518	2	Q89QJ1_BRAJA	Q89qj1 bradyrhizob
98	41	50.6	602	2	Q51T31_MAGJR	Q51t31 magnaporthe
99	41	50.6	873	2	Q64YE9_BACFR	Q64ye9 bacteroides
100	41	50.6	888	2	Q5LHH5_BACFN	Q5lhh5 bacteroides
101	41	50.6	1874	2	Q627X6_CAEBR	Q627x6 caenorhabdi
102	41	50.6	2395	2	Q6Q9H2_CAEEL	Q6q9h2 caenorhabdi
103	41	50.6	2407	2	Q9NEL2_CAEEL	Q9nel2 caenorhabdi
104	40.5	50.0	87	2	Q515N5_ENTHI	Q515n5 entamoeba h

105	40.5	50.0	719	2	Q86S17_CAEEL	Q86s17 caenorhabdi	178	39	48.1	423	2	Q5CGN4_CRYHO	Q5cgn4 cryptospori
106	40.5	50.0	733	2	Q4QFY0_LEIMA	Q4qfy0 leishmania	179	39	48.1	437	2	Q5C307_SCHJA	Q5c307 schistosoma
107	40.5	50.0	880	2	Q9TXN6_CAEEL	Q9txn6 caenorhabdi	180	39	48.1	510	1	CP191_CARAU	P79690 carassius a
108	40.5	50.0	936	2	Q624X9_CAEBR	Q624x9 caenorhabdi	181	39	48.1	512	2	Q88XS6_LACPL	Q88xs6 lactobacill
109	40	49.4	179	2	Q67B68_9CHLR	Q67b68 dehalococco	182	39	48.1	513	2	Q5FKK0_LACAC	Q5fkk0 lactobacill
110	40	49.4	184	2	Q5ZUT8_LEGPH	Q5zut8 legionella	183	39	48.1	525	2	Q19893_CAEEL	Q19893 caenorhabdi
111	40	49.4	189	2	Q6VSQ8_9PASS	Q6vsq8 muscicap a	184	39	48.1	531	2	Q60YX3_CAEBR	Q60yx3 caenorhabdi
112	40	49.4	243	2	Q7Z3J8_HUMAN	Q7z3j8 homo sapien	185	39	48.1	532	2	Q980Y2_SULSO	Q980y2 sulfolobus
113	40	49.4	271	2	Q870P8_NEUCR	Q870p8 neurospora	186	39	48.1	568	2	Q6E437_CUCME	Q6e437 cucumis mel
114	40	49.4	272	2	Q629U1_BURMA	Q629u1 burkholderi	187	39	48.1	607	2	Q82M25_STRAW	Q82m25 streptomyce
115	40	49.4	284	2	Q7RZJ6_NEUCR	Q7rzj6 neurospora	188	39	48.1	609	2	Q9GLF0_CANFA	Q9glf0 canis famil
116	40	49.4	294	2	Q8TI08_METAC	Q8ti08 methanosarc	189	39	48.1	611	1	LYAM2_CANFA	P33730 canis famil
117	40	49.4	350	2	Q4NBW3_9MICC	Q4nbw3 arthrobacte	190	39	48.1	618	2	Q7UZF4_PROMP	Q7uzf4 prochloroco
118	40	49.4	375	2	Q9XC94_KLEPN	Q9xc94 klebsiella	191	39	48.1	646	2	Q9ZBV8_STRCO	Q9zbv8 streptomyce
119	40	49.4	375	2	Q4QTM7_KLEPN	Q4qtm7 klebsiella	192	39	48.1	652	2	Q9KF83_BACHD	Q9kf83 bacillus ha
120	40	49.4	380	2	Q7DMS9_TOBAC	Q7dms9 nicotiana t	193	39	48.1	661	2	Q8K043_MOUSE	Q8k043 mus musculu
121	40	49.4	392	2	Q43326_TOBAC	Q43326 nicotiana t	194	39	48.1	728	2	Q5R1Y5_MALDO	Q5r1y5 malus domes
122	40	49.4	436	2	Q4YRZ7_PLABE	Q4yrz7 plasmodium	195	39	48.1	748	1	SORT_RAT	Q54861 rattus norv
123	40	49.4	461	1	SSL1_YEAST	Q04673 saccharomyc	196	39	48.1	812	2	Q7S5I9_NEUCR	Q7s5i9 neurospora
124	40	49.4	461	2	Q6B237_YEAST	Q6b237 saccharomyc	197	39	48.1	825	2	Q9QXW6_MOUSE	Q9qxw6 mus musculu
125	40	49.4	472	2	Q4ZRV3_PSESY	Q4zrv3 pseudomonas	198	39	48.1	825	2	Q6PHU5_MOUSE	Q6phu5 mus musculu
126	40	49.4	478	2	Q7T1K1_BRARE	Q7tlk1 brachydanio	199	39	48.1	827	2	Q5CKM9_CRYHO	Q5ckm9 cryptospori
127	40	49.4	497	1	ARAA_BACHD	Q9kbq2 bacillus ha	200	39	48.1	831	1	SORT_HUMAN	Q99523 homo sapien
128	40	49.4	500	1	ATG4_PODAN	Q86z15 podospora a	201	39	48.1	831	2	Q8I249_HUMAN	Q8i249 homo sapien
129	40	49.4	527	2	Q5WBZ7_BACSK	Q5wbz7 bacillus cl	202	39	48.1	833	2	Q8N6W5_HUMAN	Q8n6w5 homo sapien
130	40	49.4	562	2	Q869B7_LEIME	Q869b7 leishmania	203	39	48.1	834	2	Q5CYI3_CRYPV	Q5cyi3 cryptospori
131	40	49.4	562	2	Q8I911_LEIAM	Q8i911 leishmania	204	39	48.1	891	2	Q8D0K2_YERPE	Q8d0k2 yersinia pe
132	40	49.4	562	2	Q4Q3K1_LEIMA	Q4q3k1 leishmania	205	39	48.1	891	2	Q8ZEJ0_YERPE	Q8zej0 yersinia pe
133	40	49.4	569	1	RS1_CHLTR	Q84100 chlamydia t	206	39	48.1	891	2	Q66AM7_YERPS	Q66am7 yersinia ps
134	40	49.4	577	2	Q5SY54_HUMAN	Q5sy54 homo sapien	207	39	48.1	892	2	Q57NR1_SALCH	Q57nr1 salmonella
135	40	49.4	672	2	Q4H4S4_9DEIO	Q4h4s4 deinococcus	208	39	48.1	892	2	Q8Z7F1_SALTY	Q8z7f1 salmonella
136	40	49.4	688	2	Q4T0K3_TETNG	Q4t0k3 tetraodon n	209	39	48.1	892	2	Q8ZP45_SALTY	Q8zp45 salmonella
137	40	49.4	802	2	Q60YM9_CAEBR	Q60ym9 caenorhabdi	210	39	48.1	892	2	Q5PCY6_SALPA	Q5pcy6 salmonella
138	40	49.4	802	2	Q95QC4_CAEEL	Q95qc4 caenorhabdi	211	39	48.1	908	2	Q9C0G5_HUMAN	Q9c0g5 homo sapien
139	40	49.4	890	2	Q5SY53_HUMAN	Q5sy53 homo sapien	212	39	48.1	973	2	Q5VR95_ORYSA	Q5vr95 oryza sativ
140	40	49.4	890	2	Q7Z6I6_HUMAN	Q7z6i6 homo sapien	213	39	48.1	979	2	Q4RPC8_TETNG	Q4rpc8 tetraodon n
141	40	49.4	917	2	Q803S9_BRARE	Q803s9 brachydanio	214	39	48.1	1019	2	Q6P9B9_HUMAN	Q6p9b9 homo sapien
142	40	49.4	1029	2	Q7RFV6_PLAYO	Q7rfv6 plasmodium	215	39	48.1	1036	2	Q8A2G5_BACTN	Q8a2g5 bacteroides
143	40	49.4	1084	2	Q51VB3_MAGGR	Q51vb3 magnaporthe	216	39	48.1	1049	1	ADP1_YEAST	P25371 saccharomyc
144	40	49.4	1093	1	SYNP2_HUMAN	Q9ums6 homo sapien	217	39	48.1	1103	2	Q54R25_DICDI	Q54r25 dictyosteli
145	40	49.4	1110	2	Q6ZML6_HUMAN	Q6zml6 homo sapien	218	39	48.1	1190	2	Q4KG87_PSEFS	Q4kg87 pseudomonas
146	40	49.4	1209	2	Q54SA2_DICDI	Q54sa2 dictyosteli	219	39	48.1	1234	2	Q5ZM67_CHICK	Q5zm67 gallus gall
147	40	49.4	1382	2	Q9I8Y2_BRARE	Q9i8y2 brachydanio	220	39	48.1	1280	2	Q6QHS1_LYTVA	Q6qhs1 lytechinus
148	40	49.4	1520	2	Q4Q7T0_LEIMA	Q4q7t0 leishmania	221	39	48.1	1588	1	ARO1_YEAST	P08566 s pentafunc
149	40	49.4	3283	2	Q4UI21_THEAN	Q4ui21 theileria a	222	39	48.1	1638	2	Q61WT5_CAEBR	Q61wt5 caenorhabdi
150	39.5	48.8	113	2	Q8X3G6_ECO57	Q8x3g6 escherichia	223	39	48.1	1642	2	O62055_CABEL	O62055 caenorhabdi
151	39.5	48.8	113	2	Q8X5F5_ECO57	Q8x5f5 escherichia	224	38.5	47.5	512	2	Q4WSI1_TETNG	Q4sws1 tetraodon n
152	39.5	48.8	696	2	Q5CVJ1_CRYPV	Q5cvj1 cryptospori	225	38.5	47.5	665	1	KUP2_LACJO	Q5flf5 lactobacill
153	39	48.1	120	2	Q736R5_BACC1	Q736r5 bacillus ce	226	38.5	47.5	671	1	KUP2_LACAC	Q4re45 tetraodon n
154	39	48.1	134	2	Q82UU9_NITEU	Q82uu9 nitrosomona	227	38.5	47.5	746	2	Q4RE45_TETNG	Q8lfv3 bacillus ce
155	39	48.1	139	2	Q9PYE2_9BROM	Q9pye2 cucumber mo	228	38.5	47.5	859	2	Q81FV3_BACCR	Q81f73 bacillus an
156	39	48.1	161	2	Q92X31_RHIME	Q92x31 rhizobium m	229	38.5	47.5	861	2	Q6I187_BACAN	Q6i187 bacillus an
157	39	48.1	165	2	Q8TNM0_METAC	Q8tnm0 methanosarc	230	38.5	47.5	861	2	Q4MK61_BACCE	Q4mk61 bacillus ce
158	39	48.1	187	2	Q9RYS5_DEIRA	Q9rys5 deinococcus	231	38.5	47.5	861	2	Q6HL89_BACHK	Q6hl89 bacillus th
159	39	48.1	200	2	P87426_9AVES	P87426 gavia arcti	232	38.5	47.5	861	2	Q73B30_BACCI	Q73b30 bacillus ce
160	39	48.1	202	2	Q8QH35_9PASS	Q8qh35 lepidothrix	233	38.5	47.5	861	2	Q81T06_BACAN	Q81t06 bacillus an
161	39	48.1	213	2	Q9Z371_HELPJ	Q9z371 helicobacte	234	38.5	47.5	861	2	Q63DR6_BACCZ	Q63dr6 bacillus ce
162	39	48.1	220	2	Q4NI67_9MICC	Q4ni67 arthrobacte	235	38.5	47.5	901	2	Q9LR43_ARATH	Q9l43 arabidopsis
163	39	48.1	236	2	Q83267_9BROM	Q83267 cucumber mo	236	38.5	47.5	1878	2	Q9SYJ7_ARATH	Q9syj7 arabidopsis
164	39	48.1	244	2	Q91BB7_NPVST	Q91bb7 spodoptera	237	38.5	47.5	1950	2	Q9AUE0_ARATH	Q9aue0 arabidopsis
165	39	48.1	283	2	Q8SWL0_ENCCU	Q8swl0 encephalito	238	38.5	47.5	2109	2	Q5R9F7_PONPY	Q5r9f7 pongo pygma
166	39	48.1	287	2	Q51E80_ENTHI	Q51e80 entamoeba h	239	38.5	47.5	8515	2	Q4WVN4_ASPFU	Q4wvn4 aspergillus
167	39	48.1	292	2	Q5CKF5_CRYHO	Q5ckf5 cryptospori	240	38	46.9	69	2	Q4WDN2_ASPFU	Q4wdn2 aspergillus
168	39	48.1	349	2	Q4RYU8_TETNG	Q4ryu8 tetraodon n	241	38	46.9	71	2	Q4L1N0_MICAG	Q4l1n0 microtus ag
169	39	48.1	354	1	ALR_FUSNN	Q8rga2 fusobacteri	242	38	46.9	94	2	O20627_PHYPO	O20627 physarum po
170	39	48.1	361	2	Q5AXD6_EMENI	Q5axd6 aspergillus	243	38	46.9	105	2	Q9HTR9_PSEAE	Q9htr9 pseudomonas
171	39	48.1	399	2	Q7T0P4_XENLA	Q7t0p4 xenopus lae	244	38	46.9	109	2	Q88CE9_PSEPK	Q88ce9 pseudomonas
172	39	48.1	406	2	Q50XJ2_ENTHI	Q50xj2 entamoeba h	245	38	46.9	117	2	Q84258_PAPVE	Q84258 european el
173	39	48.1	409	2	Q55P67_CRYNE	Q55p67 cryptococcu	246	38	46.9	122	2	Q4YNP2_PLABE	Q4ynp2 plasmodium
174	39	48.1	409	2	Q5KE90_CRYNE	Q5ke90 cryptococcu	247	38	46.9	128	2	Q4RBS7_TETNG	Q4rbs7 tetraodon n
175	39	48.1	416	2	Q39235_9GAMA	Q39235 rhesus monk	248	38	46.9	136	2	Q752L8_ASHGO	Q752l8 ashbya goss
176	39	48.1	416	2	Q77VI8_9GAMA	Q77vi8 macaca mula	249	38	46.9	142	2	Q37060_ACACA	Q37060 acanthamoeb
177	39	48.1	416	2	Q53D52_9GAMA	Q53d52 macaca fusc	250	38	46.9	145	2	Q4RPI8_TETNG	Q4rpi8 tetraodon n

251	38	46.9	166	2	Q6ZL45_ORYSA	Q6z145	oryza sativ	324	38	46.9	775	2	Q8DJ10_SYNEL	Q8dji0	synechococc
252	38	46.9	177	2	Q626B5_CAEBR	Q626b5	caenorhabdi	325	38	46.9	840	2	Q5THT2_HUMAN	Q5tht2	homo sapien
253	38	46.9	177	2	Q8BWC2_MOUSE	Q8bwc2	mus musculus	326	38	46.9	867	2	Q4PG03_USTMA	Q4pg03	ustilago ma
254	38	46.9	198	2	Q4J5N8_AZOVI	Q4j5n8	azotobacter	327	38	46.9	877	2	Q7W5Z8_BORPA	Q7w5z8	bordetella
255	38	46.9	206	2	Q6MPH5_BDEBA	Q6mph5	bdellovibri	328	38	46.9	877	2	Q7WGUI_BORBR	Q7wgul	bordetella
256	38	46.9	207	2	Q5LA87_BACFN	Q5la87	bacteroides	329	38	46.9	893	2	Q7VZG5_BORPE	Q7vzg5	bordetella
257	38	46.9	212	2	Q87BD3_XYLFT	Q87bd3	xylella fas	330	38	46.9	952	2	Q6CEB7_YARLI	Q6ceb7	yarrowia li
258	38	46.9	213	2	Q25918_HELPY	Q25918	helicobacte	331	38	46.9	961	2	Q9SXA3_ARATH	Q9sxa3	arabidopsis
259	38	46.9	215	1	END3_BUCBP	Q89aw4	buchnera ap	332	38	46.9	1032	2	Q7YYX2_CRYPV	Q7yyx2	cryptospori
260	38	46.9	218	2	Q9PFL0_XYLFA	Q9pfl0	xylella fas	333	38	46.9	1081	2	Q7YUL8_9TRYP	Q7yul8	trypanosoma
261	38	46.9	225	2	Q5WW04_LEGPL	Q5ww04	legionella	334	38	46.9	1243	2	Q6ZRQ5_HUMAN	Q6zrq5	homo sapien
262	38	46.9	225	2	Q5X4L4_LEGPA	Q5x4l4	legionella	335	38	46.9	1672	2	Q5T4S8_HUMAN	Q5t4s8	homo sapien
263	38	46.9	225	2	Q5ZU05_LEGPH	Q5zuu5	legionella	336	38	46.9	1678	2	Q9P2N9_HUMAN	Q9p2n9	homo sapien
264	38	46.9	225	2	Q54169_STREPTOMYCE	Q54169	streptomyce	337	38	46.9	2056	2	Q8CHF3_MOUSE	Q8chf3	mus musculus
265	38	46.9	256	2	Q9S4D5_STRFR	Q9s4d5	streptomyce	338	38	46.9	2058	2	Q5T4S9_HUMAN	Q5t4s9	homo sapien
266	38	46.9	259	2	Q9RRQ0_DEIRA	Q9rrq0	deinococcus	339	38	46.9	2368	2	Q6LFN0_PLAF7	Q6lfn0	plasmodium
267	38	46.9	262	2	Q4WDH7_ASPFU	Q4wdh7	aspergillus	340	38	46.9	2553	2	Q22860_CAEEEL	Q22860	caenorhabdi
268	38	46.9	267	2	Q6N1W6_RHOPA	Q6n1w6	rhodopseudo	341	38	46.9	4643	2	Q51XM7_MAGGR	Q51xm7	magnaporthe
269	38	46.9	285	2	Q63F10_BACC2	Q63f10	bacillus ce	342	38	46.9	5129	2	Q4S817_TETNG	Q4s817	tetraodon n
270	38	46.9	285	2	Q81U89_BACAN	Q81u89	bacillus an	343	38	46.9	5183	2	Q8TDN5_HUMAN	Q8tdn5	homo sapien
271	38	46.9	295	2	Q8YPB9_ANASP	Q8ypb9	anabaena sp	344	38	46.9	5183	2	Q5T4S7_HUMAN	Q5t4s7	homo sapien
272	38	46.9	310	2	Q64924_MAIZE	Q64924	zea mays (m	345	37.5	46.3	262	2	Q9HT11_PSEAE	Q9ht11	pseudomonas
273	38	46.9	323	2	Q9KFL1_BACHD	Q9kfl1	bacillus ha	346	37.5	46.3	279	2	Q7PW18_ANOGA	Q7pw18	anopheles g
274	38	46.9	337	2	Q41975_MHV68	Q41975	murid herpe	347	37.5	46.3	327	2	Q8TFW5_ASPFU	Q8tfw5	aspergillus
275	38	46.9	342	2	Q4KL87_XENLA	Q4kl87	xenopus lae	348	37.5	46.3	327	2	Q4WSE6_ASPFU	Q4wse6	aspergillus
276	38	46.9	379	2	Q83DG6_COXBU	Q83dg6	coxiella bu	349	37.5	46.3	364	1	MURG_THETN	Q8r9g6	thermoanaer
277	38	46.9	381	2	Q5F812_NEIG1	Q5f812	neisseria g	350	37.5	46.3	485	2	Q589Y1_TOBAC	Q589y1	nicotiana t
278	38	46.9	388	2	Q7ZY44_XENLA	Q7zy44	xenopus lae	351	37.5	46.3	584	2	Q4WJ59_ASPFU	Q4wj59	aspergillus
279	38	46.9	397	2	Q971N9_SULTO	Q971n9	sulfolobus	352	37.5	46.3	683	1	KUP1_LACJO	Q74ln3	lactobacill
280	38	46.9	401	2	Q8KJ51_RHILO	Q8kj51	rhizobium l	353	37.5	46.3	756	2	Q28483_MACFA	Q28483	macaca fasc
281	38	46.9	421	1	Y118_TREPA	Q83155	treponema p	354	37.5	46.3	786	2	Q4SR20_TETNG	Q4sr20	tetraodon n
282	38	46.9	422	2	Q5KEZ8_CRYNE	Q5kez8	cryptococcu	355	37.5	46.3	829	2	Q4UBD0_THEAN	Q4ubd0	theileria a
283	38	46.9	424	2	Q98AU7_RHILO	Q98au7	rhizobium l	356	37.5	46.3	935	1	ADA22_XENLA	Q42596	xenopus lae
284	38	46.9	425	1	G6PD_HELPY	P56110	helicobacte	357	37.5	46.3	1892	2	Q51JA3_9VIRU	Q51ja3	apricot pse
285	38	46.9	440	2	Q7NJUD4_GLOVI	Q7njd4	gloeobacter	358	37	45.7	49	2	Q5C720_SCHJA	Q5c720	schistosoma
286	38	46.9	447	1	DHE4_SALTI	Q826f6	salmonella	359	37	45.7	68	2	Q7MIR3_VIBVY	Q7mir3	vibrio vuln
287	38	46.9	447	1	DHE4_SALTY	P15111	salmonella	360	37	45.7	101	2	Q9YBL4_AERPE	Q9ybl4	aeropyrum p
288	38	46.9	447	2	Q57PY4_SALCH	Q57py4	salmonella	361	37	45.7	108	2	Q5H280_XANOR	Q5h280	xanthomonas
289	38	46.9	447	2	Q5PHC4_SALPA	Q5phc4	salmonella	362	37	45.7	122	2	Q9QJT5_9RHAB	Q9qjt5	snakehead r
290	38	46.9	455	2	Q9JN57_STRFR	Q9jnn57	streptomyce	363	37	45.7	151	2	Q58EA0_XENLA	Q58ea0	xenopus lae
291	38	46.9	460	2	Q55NC9_CRYNE	Q55nc9	cryptococcu	364	37	45.7	167	2	Q90XJ2_9ACTI	Q90xj2	acipenser s
292	38	46.9	460	2	Q5KBR6_CRYNE	Q5kbr6	cryptococcu	365	37	45.7	173	2	Q7VVG1_BORPE	Q7vvgl	bordetella
293	38	46.9	460	2	Q5KBR5_CRYNE	Q5kbr5	cryptococcu	366	37	45.7	173	2	Q7WB95_BORPA	Q7wb95	bordetella
294	38	46.9	465	2	Q8L4X7_ORYSA	Q8l4x7	oryza sativ	367	37	45.7	173	2	Q7WMR3_BORBR	Q7wmr3	bordetella
295	38	46.9	469	2	Q4KCQ9_PSEF5	Q4kcq9	pseudomonas	368	37	45.7	177	2	Q67705_9VIRU	Q67705	grapevine v
296	38	46.9	469	2	Q88H21_PSEPK	Q88h21	pseudomonas	369	37	45.7	188	2	Q6DFJ3_XENLA	Q6dfj3	xenopus lae
297	38	46.9	471	2	Q9HWP9_PSEAE	Q9hwp9	pseudomonas	370	37	45.7	192	2	Q6SGI7_9BACT	Q6sgi7	uncultured
298	38	46.9	476	2	Q8ZZR7_PYRAE	Q8zzr7	pyrobaculum	371	37	45.7	192	2	Q6SHH7_9BACT	Q6shh7	uncultured
299	38	46.9	478	2	Q9REU2_STRFR	Q9reu2	streptomyce	372	37	45.7	195	2	Q7ZTE9_PAUPA	Q7zte9	pauxi pauxi
300	38	46.9	480	2	Q8F8B3_LEPIN	Q8f8b3	leptospira	373	37	45.7	195	2	Q7ZTF0_9GALL	Q7ztf0	nothocrax u
301	38	46.9	505	2	Q61PF3_CAEBR	Q61pf3	caenorhabdi	374	37	45.7	195	2	Q7ZTF1_9GALL	Q7ztf1	mitu tubero
302	38	46.9	506	2	Q6WKZ7_MOUSE	Q6wkz7	mus musculus	375	37	45.7	195	2	Q7ZTF2_CRABL	Q7ztf2	crax blumen
303	38	46.9	506	2	Q7TSK5_MOUSE	Q7tsk5	mus musculus	376	37	45.7	195	2	Q7ZTF4_OREDE	Q7ztf4	oreophasia
304	38	46.9	513	2	Q6UP84_ALCEU	Q6up84	alcaligenes	377	37	45.7	195	2	Q7ZTF5_PIPJA	Q7ztf5	pipile jacu
305	38	46.9	514	2	Q4J8W0_SULAC	Q4j8w0	sulfolobus	378	37	45.7	195	2	Q7ZTF6_9GALL	Q7ztf6	penelopina
306	38	46.9	516	2	Q89VQ1_BRAJA	Q89vq1	bradyrhizob	379	37	45.7	195	2	Q7ZTF7_9GALL	Q7ztf7	penelope ob
307	38	46.9	517	1	VNNL2_DROME	Q8irr1	drosophila	380	37	45.7	195	2	Q7ZTF8_9GALL	Q7ztf8	chamaepetes
308	38	46.9	519	2	Q84FW2_METEX	Q84fw2	methylobact	381	37	45.7	195	2	Q7ZTF9_9GALL	Q7ztf9	aburria abu
309	38	46.9	526	2	Q7WMR8_BORBR	Q7wmr8	bordetella	382	37	45.7	196	2	Q58X02_ORENI	Q58x02	oreochromis
310	38	46.9	532	2	Q55R72_CRYNE	Q55r72	cryptococcu	383	37	45.7	198	2	Q5BPI1_ARATH	Q5bpi1	arabidopsis
311	38	46.9	543	2	Q65114_ASF	Q65114	african swi	384	37	45.7	200	2	Q90XW3_SCOUM	Q90xw3	scopus umbr
312	38	46.9	582	2	Q6C659_YARLI	Q6c659	yarrowia li	385	37	45.7	202	2	Q8QH29_9PASS	Q8qh29	ptilonorhyn
313	38	46.9	582	2	Q52ZA2_CARAU	Q52za2	carassius a	386	37	45.7	209	1	END3_BUCAI	Q8kal6	buchnera ap
314	38	46.9	616	2	Q7NGP4_GLOVI	Q7ngp4	gloeobacter	387	37	45.7	210	1	END3_BUCAI	P57219	buchnera ap
315	38	46.9	626	2	Q96JY1_HUMAN	Q96jy1	homo sapien	388	37	45.7	211	2	Q6H8Q4_CANFA	Q6h8q4	canis famil
316	38	46.9	648	2	Q8DKA9_SYNEL	Q8dka9	synechococc	389	37	45.7	214	2	Q9RYS3_DEIRA	Q9rys3	deinococcus
317	38	46.9	670	2	Q8YMH3_ANASP	Q8ymh3	anabaena sp	390	37	45.7	218	2	Q6ZW87_HUMAN	Q6zw87	homo sapien
318	38	46.9	679	2	Q5RCY5_PONPY	Q5rcy5	pongo pygma	391	37	45.7	220	2	Q8BTF9_MOUSE	Q8btf9	mus musculus
319	38	46.9	681	1	Y005_SYNV3	Q55680	synechocyst	392	37	45.7	221	2	Q4S895_TETNG	Q4s895	tetraodon n
320	38	46.9	683	2	Q6MF18_PARUW	Q6mf18	parachlamyd	393	37	45.7	223	2	Q801N6_XENLA	Q801n6	xenopus lae
321	38	46.9	718	2	Q5CJJ3_CRYHO	Q5cjj3	cryptospori	394	37	45.7	227	2	Q51EAG_ENTHI	Q51ea6	entamoeba h
322	38	46.9	719	2	Q7YYQ4_CRYPV	Q7yyq4	cryptospori	395	37	45.7	238	2	Q65QM8_MANSM	Q65qm8	mannheimia
323	38	46.9	729	2	Q8LMU8_ORYSA	Q8lmu8	oryza sativ	396	37	45.7	251	2	Q8SQF1_NATST	Q8sqf1	natalus str

397	37	45.7	255	2	Q9A2U9_CAUCR	Q9a2u9 caulobacter	470	37	45.7	541	1	CGT_MOUSE	Q64676 mus musculu
398	37	45.7	258	2	Q88QC0_PSEPK	Q88qc0 pseudomonas	471	37	45.7	541	1	CGT_RAT	Q09426 rattus norv
399	37	45.7	266	2	Q5XGL8_XENLA	Q5xgl8 xenopus lae	472	37	45.7	541	2	Q91W57_MOUSE	Q91w57 mus musculu
400	37	45.7	267	2	Q7MYV2_PHOLL	Q7myv2 photorhabdu	473	37	45.7	541	2	Q98TB5_CHICK	Q98tb5 gallus gall
401	37	45.7	272	1	PANB_XYLFA	Q9pgr9 xylella fas	474	37	45.7	543	2	Q801U7_BRARE	Q801u7 brachydanio
402	37	45.7	272	1	PANB_XYLFT	Q87ew0 xylella fas	475	37	45.7	569	2	Q6BG99_PARTE	Q6bg99 paramecium
403	37	45.7	272	2	Q6JQQ4_BOMOR	Q6jqq4 bombina ori	476	37	45.7	573	2	Q5RKI3_RAT	Q5rki3 rattus norv
404	37	45.7	291	2	Q8N1C2_HUMAN	Q8n1c2 homo sapien	477	37	45.7	586	2	Q4SU67_TETNG	Q4su67 tetraodon n
405	37	45.7	291	2	Q86SE5_HUMAN	Q8n1c2 homo sapien	478	37	45.7	593	2	Q4WAG8_ASFPU	Q4wag8 aspergillus
406	37	45.7	293	2	Q8BTG1_MOUSE	Q8btg1 mus musculu	479	37	45.7	595	2	Q414U4_GIBZE	Q414u4 gibberella
407	37	45.7	293	2	Q8BTF8_MOUSE	Q8btf8 mus musculu	480	37	45.7	595	2	Q7PHL5_ANOGA	Q7phl5 anopheles g
408	37	45.7	302	2	Q5LQ50_SILPO	Q5lq50 silicibacte	481	37	45.7	602	2	Q4S9P4_TETNG	Q4s9p4 tetraodon n
409	37	45.7	305	2	Q5M905_XENTR	Q5m905 xenopus tro	482	37	45.7	610	1	LYAM2_HORSE	Q951g1 equus cabal
410	37	45.7	308	2	Q8M77_PSEPK	Q8m77 pseudomonas	483	37	45.7	629	2	Q7Q9C3_ANOGA	Q7q9c3 anopheles g
411	37	45.7	310	2	Q98AA7_RHILO	Q98aa7 rhizobium l	484	37	45.7	638	2	Q5RDT7_PONPY	Q5rdt7 pongo pygma
412	37	45.7	314	2	Q21849_CAEEL	Q21849 caenorhabdi	485	37	45.7	645	2	Q4IUU8_AZOVI	Q4iuu8 azotobacter
413	37	45.7	318	2	Q6CLY9_KLULA	Q6cly9 kluyveromyc	486	37	45.7	670	2	Q6K2I3_ORYSA	Q6k2i3 oryza sativ
414	37	45.7	318	2	Q4SH80_TETNG	Q4sh80 tetraodon n	487	37	45.7	672	2	Q4RL61_TETNG	Q4rl61 tetraodon n
415	37	45.7	321	2	Q4L9I3_STAHI	Q4l9i3 staphylococ	488	37	45.7	677	2	Q8A3Q5_BACTN	Q8a3q5 bacteroides
416	37	45.7	328	2	Q4HLM7_CAMLA	Q4hlm7 campylobact	489	37	45.7	680	2	Q4RLZ5_TETNG	Q4rlz5 tetraodon n
417	37	45.7	329	2	Q6TBR4_TOXGO	Q6tbr4 toxoplasma	490	37	45.7	688	2	Q4RLZ5_TETNG	Q4rlz5 tetraodon n
418	37	45.7	332	2	Q86UT8_HUMAN	Q86ut8 homo sapien	491	37	45.7	688	2	Q6PBB2_BRARE	Q6pbb2 brachydanio
419	37	45.7	333	2	Q7ZWH8_BRARE	Q7zwh8 brachydanio	492	37	45.7	702	2	Q4SH09_TETNG	Q4sh09 tetraodon n
420	37	45.7	333	2	Q4S564_TETNG	Q4s564 tetraodon n	493	37	45.7	716	1	YDOA_SCHPO	O13730 schizosacch
421	37	45.7	333	2	Q98TS3_BRARE	Q98ts3 brachydanio	494	37	45.7	730	2	Q5RFV8_BRARE	Q5rfv8 brachydanio
422	37	45.7	334	2	Q4UI01_THEAN	Q4ui01 theileria a	495	37	45.7	730	2	Q5BKX4_BRARE	Q5bkx4 brachydanio
423	37	45.7	334	2	Q4N7L0_THEPA	Q4n7l0 theileria p	496	37	45.7	749	1	APBA2_PONPY	Q5bxx4 pongo pygma
424	37	45.7	338	2	Q6H9N9_YARLI	Q6h9n9 yarrowia li	497	37	45.7	750	1	APBA2_RAT	O35431 rattus norv
425	37	45.7	350	2	Q8EW16_MYCPE	Q8ew16 mycoplasma	498	37	45.7	776	1	CHS1_CANAL	P23316 candida alb
426	37	45.7	358	2	Q5FP01_GLUOX	Q5fp01 gluconobact	499	37	45.7	833	1	CWH41_YEAST	P53008 saccharomyc
427	37	45.7	370	2	Q5NG03_FRATT	Q5ng03 francisella	500	37	45.7	835	2	Q4I3L3_GIBZE	Q4i3l3 gibberella
428	37	45.7	372	2	Q6Z786_ORYSA	Q6z786 oryza sativ	501	37	45.7	856	2	Q6C0S9_YARLI	Q6c0s9 yarrowia li
429	37	45.7	376	2	Q5ER65_MONAL	Q5er65 monopterus	502	37	45.7	881	2	Q8JAG5_SIVCZ	Q8jag5 chimpanzee
430	37	45.7	387	1	ILF2_BRARE	Q6nz06 brachydanio	503	37	45.7	882	2	Q6N313_RHOPA	Q6n313 rhodopseudo
431	37	45.7	408	2	Q53JE9_ORYSA	Q53je9 oryza sativ	504	37	45.7	931	2	Q5E319_VIBF1	Q5e319 vibrio fisc
432	37	45.7	413	2	Q7YRC5_BOVIN	Q7yrc5 bos taurus	505	37	45.7	946	2	Q57V94_9TRYP	Q57v94 trypanosoma
433	37	45.7	414	2	Q5N1M6_SYN6	Q5nlm6 synchococc	506	37	45.7	980	2	Q5U167_DROME	Q5u167 drosophila
434	37	45.7	415	2	Q6INN9_XENLA	Q6inn9 xenopus lae	507	37	45.7	1005	2	Q5K995_CRYNE	Q5k995 cryptococcu
435	37	45.7	416	2	Q9XX64_CAEEL	Q9xx64 caenorhabdi	508	37	45.7	1005	2	Q55KA7_CRYNE	Q55ka7 cryptococcu
436	37	45.7	424	2	Q5NIN0_FRATT	Q5nin0 francisella	509	37	45.7	1005	2	Q6BZ46_DEBHA	Q6bz46 debaryomyce
437	37	45.7	434	2	Q4RY62_TETNG	Q4ry62 tetraodon n	510	37	45.7	1012	2	Q6BZ46_DEBHA	Q6bz46 debaryomyce
438	37	45.7	436	2	Q7UVS3_RHOPA	Q7uvs3 rhodopirell	511	37	45.7	1022	2	Q6CS24_KLULA	Q6cs24 kluyveromyc
439	37	45.7	438	2	Q6SH08_9BACT	Q6sh08 uncultured	512	37	45.7	1026	2	Q5A594_CANAL	Q5a594 candida alb
440	37	45.7	444	2	Q51NM2_MAGGR	Q51nm2 magnaporth	513	37	45.7	1026	2	Q9URM1_CANAL	Q9urm1 candida alb
441	37	45.7	447	2	Q5P0F8_AZOSE	Q5p0f8 azoarcus sp	514	37	45.7	1044	2	Q8QUS6_9VIRU	Q8qus6 infectious
442	37	45.7	454	2	Q864N3_NATST	Q864n3 natalus str	515	37	45.7	1045	2	Q8BM83_MOUSE	Q8bm83 m mus muscu
443	37	45.7	454	2	Q53AI6_9CHIR	Q53ai6 natalus tum	516	37	45.7	1052	2	Q4KSC1_9VIRU	Q4ksc1 orange-spot
444	37	45.7	454	2	Q53AI7_9CHIR	Q53ai7 natalus sat	517	37	45.7	1053	2	Q5YF54_9VIRU	Q5yf54 rock bream
445	37	45.7	454	2	Q53AI8_9CHIR	Q53ai8 natalus jam	518	37	45.7	1055	2	Q6FKY0_CANGA	Q6fky0 candida gla
446	37	45.7	454	2	Q53AI9_9CHIR	Q53ai9 natalus maj	519	37	45.7	1105	2	Q55FS2_DICDI	Q55fs2 dictyosteli
447	37	45.7	454	2	Q53AJ0_9CHIR	Q53aj0 chilonatalu	520	37	45.7	1106	2	Q42291_CHICK	Q42291 gallus gall
448	37	45.7	454	2	Q53AJ1_9CHIR	Q53aj1 nyctiellus	521	37	45.7	1142	2	Q26615_STRPU	Q26615 strongyloce
449	37	45.7	455	2	Q7VZQ0_BORPE	Q7vzq0 bordetella	522	37	45.7	1168	2	Q60XC0_CAEER	Q60xc0 caenorhabdi
450	37	45.7	455	2	Q7W5B2_BORPA	Q7w5b2 bordetella	523	37	45.7	1212	2	Q4RL28_TETNG	Q4rl28 tetraodon n
451	37	45.7	455	2	Q7WCU4_BORBR	Q7wcu4 bordetella	524	37	45.7	1244	1	MIX1_CAEEL	Q09591 caenorhabdi
452	37	45.7	461	2	Q86KZ2_DICDI	Q86kz2 dictyosteli	525	37	45.7	1279	2	Q86PE6_DROME	Q86pe6 drosophila
453	37	45.7	468	2	Q52KU4_XENLA	Q52ku4 xenopus lae	526	37	45.7	1296	2	Q61ZQ5_CAEER	Q61zg5 caenorhabdi
454	37	45.7	469	1	GLNA_PROVU	p28786 proteus vul	527	37	45.7	1307	2	Q9GT78_PLAFA	Q9gt78 plasmodium
455	37	45.7	473	2	Q9LJ71_ARATH	Q9lj71 arabidopsis	528	37	45.7	1322	2	Q7X7E9_ORYSA	Q7x7e9 oryza sativ
456	37	45.7	473	2	Q7WBH7_BORPA	Q7wbh7 bordetella	529	37	45.7	1368	2	Q8I199_PLAF7	Q8i199 plasmodium
457	37	45.7	473	2	Q7WMZ9_BORBR	Q7wmz9 bordetella	530	37	45.7	1411	2	Q5ARI1_EMENI	Q5ari1 aspergillus
458	37	45.7	478	2	Q74H74_GEOSL	Q74h74 geobacter s	531	37	45.7	1460	2	Q54HC6_DICDI	Q54hc6 dictyosteli
459	37	45.7	489	2	Q569D6_XENTR	Q569d6 xenopus tro	532	37	45.7	1479	2	Q5BFD0_EMENI	Q5bfd0 aspergillus
460	37	45.7	490	2	Q8RZL9_ORYSA	Q8rzl9 oryza sativ	533	37	45.7	1544	2	Q9W2F2_DROME	Q9w2f2 drosophila
461	37	45.7	491	2	Q62FI2_BURMA	Q62fi2 burkholderi	534	37	45.7	1664	2	Q9TVQ2_CAEEL	Q9tvq2 caenorhabdi
462	37	45.7	491	2	Q63VC1_BURPS	Q63vc1 burkholderi	535	37	45.7	1912	1	CHD4_HUMAN	Q14839 homo sapien
463	37	45.7	493	2	Q65GC0_BACLD	Q65gc0 bacillus li	536	37	45.7	1915	1	CHD4_MOUSE	Q6pdk2 mus musculu
464	37	45.7	496	2	Q6C7M6_YARLI	Q6c7m6 yarrowia li	537	37	45.7	1945	2	Q5DTP7_MOUSE	Q5dtp7 mus musculu
465	37	45.7	498	2	Q9BDZ8_BOVIN	Q9bdz8 bos taurus	538	37	45.7	2248	2	Q4S447_TETNG	Q4s447 tetraodon n
466	37	45.7	505	2	Q4SSL0_TETNG	Q4ssl0 tetraodon n	539	37	45.7	2386	1	RAD3_SCHPO	Q02099 schizosacch
467	37	45.7	518	2	Q6NBU5_RHOPA	Q6nbu5 rhodopseudo	540	37	45.7	2529	2	Q7QJP5_ANOGA	Q7qjp5 anopheles g
468	37	45.7	530	2	Q8NSQ1_CORGL	Q8nsq1 corynebacte	541	37	45.7	2531	1	NOTC1_RAT	Q07008 rattus norv
469	37	45.7	540	2	Q9E201_9ALPH	Q9e201 cercopithec	542	37	45.7	2820	1	NP1_RAT	P97526 rattus norv

689	36	44.4	459	2	Q4U3V5_CRYPA	Q4u3v5	cryphonectr
690	36	44.4	460	2	Q9AUV2_ORYSA	Q9auv2	oryza sativ
691	36	44.4	462	2	Q9AUU9_ORYSA	Q9auu9	oryza sativ
692	36	44.4	466	2	Q6FB37_ACIAD	Q6fb37	acinetobact
693	36	44.4	481	2	Q67B63_9CHLR	Q67b63	dehalococco
694	36	44.4	482	2	Q67B92_9CHLR	Q67b92	dehalococco
695	36	44.4	490	2	Q7V4I3_PROMM	Q7v4i3	prochloroco
696	36	44.4	497	2	Q6Z1A7_ORYSA	Q6z1a7	oryza sativ
697	36	44.4	497	2	Q9FYU7_BRANA	Q9fyu7	brassica na
698	36	44.4	497	2	Q52TA3_ENTAG	Q52ta3	enterobacte
699	36	44.4	505	2	Q9VZJ4_DROME	Q9vzj4	drosophila
700	36	44.4	505	2	Q8BX48_MOUSE	Q8bx48	mus musculu
701	36	44.4	510	2	Q4IMA4_GIBZE	Q4ima4	gibberella
702	36	44.4	513	2	Q9D1I8_MOUSE	Q9dl18	mus musculu
703	36	44.4	516	2	Q7XS71_ORYSA	Q7xs71	oryza sativ
704	36	44.4	517	2	Q9VJ45_DROME	Q9vj45	drosophila
705	36	44.4	522	1	COR1_SCHPO	O13688	schizosacch
706	36	44.4	527	2	Q52F89_MAGGR	Q52f89	magnaporthe
707	36	44.4	528	2	Q5B6G5_EMENI	Q5b6g5	aspergillus
708	36	44.4	530	2	O16276_CAEEL	O16276	caenorhabdi
709	36	44.4	531	2	Q516K3_ENTHI	Q516k3	entamoeba h
710	36	44.4	531	2	Q21603_CAEEL	Q21603	caenorhabdi
711	36	44.4	534	2	Q8PI72_XANAC	Q8pi72	xanthomonas
712	36	44.4	535	2	O18009_CAEEL	O18009	caenorhabdi
713	36	44.4	536	2	Q74BY3_GEOSL	Q74by3	geobacter s
714	36	44.4	537	2	Q8DIX3_SYNEL	Q8dix3	synechococc
715	36	44.4	537	2	Q5I9I6_BRARE	Q5i9i6	brachydanio
716	36	44.4	538	2	Q4FM39_9RICK	Q4fm39	candidatus
717	36	44.4	541	1	CGT_HUMAN	Q16880	homo sapien
718	36	44.4	542	1	PYRG_LEPIC	Q72s46	leptospira
719	36	44.4	542	1	PYRG_LEPIN	Q8f3j3	leptospira
720	36	44.4	542	1	PYRG_RHILO	Q98mf0	rhizobium l
721	36	44.4	542	2	Q7PYJ9_ANOGA	Q7pyj9	anopheles g
722	36	44.4	543	2	Q4P8Y4_USTMA	Q4p8y4	ustilago ma
723	36	44.4	543	2	Q6N5T4_RHOPA	Q6n5t4	rhodopseudo
724	36	44.4	543	2	Q89KU5_BRAJA	Q89ku5	bradyrhizob
725	36	44.4	548	2	Q7VQT6_CANBF	Q7vqt6	candidatus
726	36	44.4	549	2	Q9RN79_STRAT	Q9rn79	streptomyce
727	36	44.4	555	2	Q8ZF10_YERPE	Q8zff0	yersinia pe
728	36	44.4	555	2	Q669X4_YERPS	Q669x4	yersinia ps
729	36	44.4	563	2	Q8NG04_HUMAN	Q8ng04	homo sapien
730	36	44.4	564	2	Q5HR52_STAEO	Q5hr52	staphylococ
731	36	44.4	564	2	Q8CQ28_STAEP	Q8cq28	staphylococ
732	36	44.4	569	2	Q4SYK5_TETNG	Q4syk5	tetraodon n
733	36	44.4	570	2	Q5AEF5_CANAL	Q5aef5	candida alb
734	36	44.4	572	1	TMPS7_HUMAN	Q7rty8	homo sapien
735	36	44.4	572	1	TMPS7_MOUSE	Q8bik6	mus musculu
736	36	44.4	573	1	DPOL_MOUSE	Q9qxe2	mus musculu
737	36	44.4	577	2	Q967G8_NIPBR	Q967g8	nippostrong
738	36	44.4	578	2	Q4P096_USTMA	Q4p096	ustilago ma
739	36	44.4	593	2	Q84H36_9BURK	Q84h36	comamonas s
740	36	44.4	595	2	Q4LU16_9BURK	Q4lu16	burkholderi
741	36	44.4	595	2	Q62MN6_BURMA	Q62mn6	burkholderi
742	36	44.4	595	2	Q63X89_BURPS	Q63x89	burkholderi
743	36	44.4	601	1	MBHL_AZOCH	P18191	azotobacter
744	36	44.4	601	2	Q4IS63_AZOVI	Q4is63	azotobacter
745	36	44.4	602	1	MBHL_AZOVI	P21949	azotobacter
746	36	44.4	602	1	XSC_ALCXX	Q84h41	alcaligenes
747	36	44.4	602	2	Q87IY6_NEUCR	Q87ly6	neurospora
748	36	44.4	602	2	Q4IUQ0_AZOVI	Q4iuq0	azotobacter
749	36	44.4	618	2	Q55FC8_DICDI	Q55fc8	dictyosteli
750	36	44.4	619	2	Q6ZMW9_HUMAN	Q6zmw9	homo sapien
751	36	44.4	621	2	Q9H9Y1_HUMAN	Q9h9y1	homo sapien
752	36	44.4	623	2	Q4RMH4_TETNG	Q4rmh4	tetraodon n
753	36	44.4	634	2	Q517K1_ENTHI	Q517k1	entamoeba h
754	36	44.4	636	2	Q22040_CAEEL	Q22040	caenorhabdi
755	36	44.4	636	2	Q9CAR2_ARATH	Q9car2	arabidopsis
756	36	44.4	642	2	Q99P70_MOUSE	Q99p70	mus musculu
757	36	44.4	642	2	Q9D0F1_MOUSE	Q9d0f1	mus musculu
758	36	44.4	646	2	Q4WAQ3_ASPFU	Q4waq3	aspergillus
759	36	44.4	648	2	Q5LH44_BACFN	Q5lh44	bacteroides
760	36	44.4	648	2	Q64Y02_BACFR	Q64y02	bacteroides
761	36	44.4	651	2	Q70YK6_ORYSA	Q70yk6	oryza sativ

36	44.4	651	2	Q712D5_ORYSA	Q712d5	oryza sativ
36	44.4	651	2	Q852P4_ORYSA	Q852p4	oryza sativ
36	44.4	654	2	Q8BVF4_MOUSE	Q8bvf4	mus musculu
36	44.4	661	2	Q512S1_ENTHI	Q512s1	entamoeba h
36	44.4	662	2	Q7S3U9_NEUCR	Q7s3u9	neurospora
36	44.4	664	2	Q9GRV2_CAEEL	Q9grv2	caenorhabdi
36	44.4	667	2	Q4NBX4_9MICC	Q4nbx4	arthrobacte
36	44.4	674	2	Q4HY53_GIBZE	Q4hy53	gibberella
36	44.4	679	2	Q4KF95_BACHD	Q4kf95	bacillus ha
36	44.4	684	2	Q4G0S6_HUMAN	Q4g0s6	homo sapien
36	44.4	721	2	Q7S7G6_NEUCR	Q7s7g6	neurospora
36	44.4	745	2	Q4FPP5_9RICK	Q4fpp5	candidatus
36	44.4	751	2	Q580F4_TRYP	Q580f4	trypanosoma
36	44.4	764	2	Q7R0K3_GIALA	Q7r0k3	giardia lam
36	44.4	764	2	P93756_ARATH	P93756	arabidopsis
36	44.4	767	2	Q7S1F1_NEUCR	Q7s1f1	neurospora
36	44.4	777	2	Q4RL36_TETNG	Q4rl36	tetraodon n
36	44.4	780	2	Q9FVW7_ARATH	Q9fvw7	arabidopsis
36	44.4	786	2	Q6BZ28_DEBHA	Q6bz28	debaryomyce
36	44.4	796	2	O17532_CAEEL	O17532	caenorhabdi
36	44.4	803	2	Q9CB85_MYCLE	Q9cb85	mycobacteri
36	44.4	818	2	Q6CNQ6_KLULA	Q6cnq6	kluyveromyc
36	44.4	820	2	Q8GX69_ARATH	Q8gx69	arabidopsis
36	44.4	820	2	Q4HGH7_CAMCO	Q4hgh7	campylobact
36	44.4	821	2	Q5CYP1_CRYPV	Q5cyp1	cryptospori
36	44.4	821	2	Q5CP41_CRYHO	Q5cp41	cryptospori
36	44.4	826	2	P72351_MYCLE	P72351	mycobacteri
36	44.4	826	2	Q5YJ18_9HERP	Q5yj18	fibropapill
36	44.4	826	2	Q5YJ19_9HERP	Q5yj19	fibropapill
36	44.4	828	2	Q5YJ17_9HERP	Q5yj17	fibropapill
36	44.4	828	2	Q5YJ22_9HERP	Q5yj22	fibropapill
36	44.4	828	2	Q5YJ23_9HERP	Q5yj23	fibropapill
36	44.4	828	2	Q5YJ24_9HERP	Q5yj24	fibropapill
36	44.4	839	2	Q9LFF3_ARATH	Q9lff3	arabidopsis
36	44.4	852	2	Q5ZR69_9HERP	Q5zr69	hawaiian gr
36	44.4	854	2	Q5Y967_9HERP	Q5y967	fibropapill
36	44.4	864	2	Q95QI8_CAEEL	Q95qi8	caenorhabdi
36	44.4	866	2	Q8EF55_SHEON	Q8ef55	shewanella
36	44.4	877	2	Q7N458_PHOLL	Q7n458	photorhabdu
36	44.4	886	2	Q8JAH4_SIVCZ	Q8jah4	chimpanzee
36	44.4	887	2	Q7NYY5_CHRVO	Q7nyy5	chromobacte
36	44.4	890	1	ADHE_ECO57	P0a9q8	escherichia
36	44.4	890	1	ADHE_ECOLI	P0a9q7	escherichia
36	44.4	891	2	Q7DLI8_ECOLI	Q7dl18	escherichia
36	44.4	891	2	Q8FHX2_ECOL6	Q8fhx2	escherichia
36	44.4	891	2	Q6D4R4_ERWCT	Q6d4r4	erwinia car
36	44.4	891	2	Q83RN2_SHIFL	Q83rn2	shigella fl
36	44.4	892	2	Q95QI7_CAEEL	Q95qi7	caenorhabdi
36	44.4	897	2	Q6VG43_SIVCZ	Q6vg43	chimpanzee
36	44.4	928	1	VGLB_BHV1P	P17471	bovine herp
36	44.4	932	1	VGLB_BHV1C	P12640	bovine herp
36	44.4	938	2	Q580W5_9TRYP	Q580w5	trypanosoma
36	44.4	938	2	Q51BQ2_ENTHI	Q51bq2	entamoeba h
36	44.4	938	2	Q722D3_LISMF	Q722d3	listeria mo
36	44.4	938	2	Q8Y8V7_LISMO	Q8y8v7	listeria mo
36	44.4	938	2	Q92DN5_LISIN	Q92dn5	listeria in
36	44.4	943	2	Q4IMI9_GIBZE	Q4imi9	gibberella
36	44.4	959	2	Q6Z572_ORYSA	Q6z572	oryza sativ
36	44.4	1022	2	Q4UHY6_THEAN	Q4uhy6	theileria a
36	44.4	1027	2	Q9BWX2_HUMAN	Q9bwx2	homo sapien
36	44.4	1027	2	Q86VS3_HUMAN	Q86vs3	homo sapien
36	44.4	1032	2	Q4N7M3_THEPA	Q4n7m3	theileria p
36	44.4	1035	2	Q7YSI2_CAEEL	Q7ysi2	caenorhabdi
36	44.4	1053	2	Q96KJ4_HUMAN	Q96kj4	homo sapien
36	44.4	1056	1	K125_ARATH	P82266	arabidopsis
36	44.4	1074	2	Q45878_CAEEL	Q45878	caenorhabdi
36	44.4	1093	2	Q640N3_MOUSE	Q640n3	mus musculu
36	44.4	1102	2	Q55WU4_CRYNE	Q55wu4	cryptococcu
36	44.4	1102	2	Q5KJF4_CRYNE	Q5kjf4	cryptococcu
36	44.4	1156	2	O61137_LEIDO	O61137	leishmania
36	44.4	1157	2	Q4Q2M2_LEIMA	Q4q2m2	leishmania
36	44.4	1161	2	Q8J213_KLULA	Q8j213	kluyveromyc

835	1235	44.4	1	RENT1_ARATH	Q9fjr0 arabidopsis	908	35	43.2	108	2	Q500B2_PSESY	Q500b2 pseudomonas
836	1243	44.4	2	Q8S3K7_ARATH	Q8s3k7 arabidopsis	909	35	43.2	109	2	Q8RPA8_PSEST	Q8rpa8 pseudomonas
837	1310	44.4	2	Q4Q2H5_LEIMA	Q4q2h5 leishmania	910	35	43.2	109	2	Q88B13_PSESM	Q88b13 pseudomonas
838	1318	44.4	2	Q6ZUA9_HUMAN	Q6zua9 homo sapien	911	35	43.2	113	2	Q9BIF1_9BILA	Q9bif1 cooperia pu
839	1341	44.4	2	Q9U1E5_LEIMA	Q9ule5 leishmania	912	35	43.2	113	2	Q9BIF2_9BILA	Q9bif2 cooperia pu
840	1387	44.4	2	Q6L1T5_CAEBR	Q6llt5 caenorhabdi	913	35	43.2	113	2	Q9BIF3_9BILA	Q9bif3 cooperia pu
841	1392	44.4	2	Q6LF61_PLAF7	Q6lf61 plasmodium	914	35	43.2	123	2	Q881M8_PSESM	Q881m8 pseudomonas
842	1491	44.4	2	Q4ID49_GIBZE	Q4id49 gibberella	915	35	43.2	124	2	Q5BXC4_SCHJA	Q5bxc4 schistosoma
843	1977	44.4	2	Q5ATT0_EMENI	Q5att0 aspergillus	916	35	43.2	125	2	Q5MK13_9ASPA	Q5mk13 lycoris sp.
844	2848	44.4	2	Q7RY33_NEUCR	Q7ry33 neurospora	917	35	43.2	127	2	Q9TV08_CANFA	Q9tv08 canis famil
845	3570	44.4	2	Q7Q737_ANOGA	Q7q737 anopheles g	918	35	43.2	128	2	Q8HY59_DEBHA	Q8hy59 oryctolagus
846	4144	43.8	88	Q8WN22_CANFA	Q8wn22 canis famil	919	35	43.2	131	2	Q6BNN2_DEBHA	Q6bnn2 debaryomyce
847	35.5	43.8	2	Q9UKK0_HUMAN	Q9ukk0 homo sapien	920	35	43.2	131	2	Q8VTN8_HELPY	Q8vtn8 helicobacte
848	119	43.8	2	Q5CCY7_9GEMI	Q5ccy7 tomato leaf	921	35	43.2	132	2	Q4HAD6_9DEIO	Q4had6 deinococcus
849	123	43.8	2	Q5CCZ1_9GEMI	Q5ccz1 tomato leaf	922	35	43.2	140	2	Q8VTN4_HELPY	Q8vtn4 helicobacte
850	158	43.8	2	Q8C9S9_MOUSE	Q8c9s9 mus musculu	923	35	43.2	140	2	Q8VTP4_HELPY	Q8vtp4 helicobacte
851	159	43.8	2	Q4RV20_TETNG	Q4rv20 tetraodon n	924	35	43.2	149	1	Y2595_SULSO	Q97vm6 sulfolobus
852	165	43.8	2	Q58GM4_9HIV1	Q58gm4 human immun	925	35	43.2	149	2	Q9VWX7_DROME	Q9vwx7 drosophila
853	204	43.8	2	Q54HQ8_DICDI	Q54hq8 dictyosteli	926	35	43.2	149	2	Q4RSP6_TETNG	Q4rsp6 tetraodon n
854	257	43.8	2	Q75QI6_9GEMI	Q75qi6 tomato leaf	927	35	43.2	150	2	Q9ZY20_9MOLL	Q9zy20 rossia palp
855	257	43.8	2	Q80S73_9GEMI	Q80s73 tomato leaf	928	35	43.2	160	2	Q7U6B3_SYNXP	Q7u6b3 synechococc
856	263	43.8	2	Q4ZL15_PSESY	Q4zll5 pseudomonas	929	35	43.2	160	2	Q5N479_SYNXP	Q5n479 synechococc
857	263	43.8	2	Q87TS5_PSESM	Q87ts5 pseudomonas	930	35	43.2	162	2	Q5MK12_9ASPA	Q5mk12 lycoris sp.
858	264	43.8	2	Q8N3V6_HUMAN	Q8n3v6 homo sapien	931	35	43.2	166	2	Q4QA21_LEIMA	Q4qa21 leishmania
859	265	43.8	2	Q898C6_CLOTE	Q898c6 clostridium	932	35	43.2	166	2	Q9N9V6_LEIIN	Q9n9v6 leishmania
860	265	43.8	2	Q8C976_MOUSE	Q8c976 mus musculu	933	35	43.2	166	2	Q8VLE2_HELPY	Q8vle2 helicobacte
861	273	43.8	2	Q4RFL4_TETNG	Q4rfl4 tetraodon n	934	35	43.2	167	2	Q5VWQ3_HUMAN	Q5vwq3 homo sapien
862	311	43.8	2	Q9VLW8_DROME	Q9vlw8 drosophila	935	35	43.2	167	2	Q4HJL7_CAMLA	Q4hj17 campylobact
863	334	43.8	1	GBLP_ORYSA	P49027 oryza sativ	936	35	43.2	168	2	Q50VB7_ENTHI	Q50vb7 entamoeba h
864	339	43.8	2	Q5AFI5_CANAL	Q5afi5 candida alb	937	35	43.2	170	2	Q51WS9_MAGGR	Q51ws9 magnaportha
865	342	43.8	2	Q5XLY0_GINBI	Q5xly0 ginkgo bilo	938	35	43.2	177	2	Q4T246_TETNG	Q4t246 tetraodon n
866	360	43.8	2	Q5SMK6_ORYSA	Q5smk6 oryza sativ	939	35	43.2	184	2	Q4HT57_CAMUP	Q4ht57 campylobact
867	363	43.8	1	MURG_LISIN	Q929Y2 listeria in	940	35	43.2	185	2	Q7P056_CHRVO	Q7p056 chromobacte
868	424	43.8	2	Q8IMS3_DROME	Q8ims3 drosophila	941	35	43.2	186	1	KAD_TROW8	Q83i60 tropheryma
869	439	43.8	2	Q5BK6G_XENTR	Q5bk6g xenopus tro	942	35	43.2	186	1	KAD_TROWT	Q83g05 tropheryma
870	567	43.8	2	Q4R6Q6_MACFA	Q4r6q6 macaca fasc	943	35	43.2	187	2	Q9A6H2_CAUCR	Q9a6h2 caulobacter
871	609	43.8	2	O04208_ARATH	Q04208 arabidopsis	944	35	43.2	189	2	Q6VSP7_9PASS	Q6vsp7 thamnolea c
872	624	43.8	2	Q8SZ15_DROME	Q8sz15 drosophila	945	35	43.2	191	2	Q65QQ8_MANSM	Q65qq8 mannheimia
873	650	43.8	1	KUP1_LACAC	Q5fmk7 lactobacill	946	35	43.2	192	2	Q5SDB2_VIBAN	Q5sdb2 vibrio angu
874	682	43.8	2	Q9STF4_ARATH	Q9stf4 arabidopsis	947	35	43.2	192	2	Q73IS9_WOLPM	Q73is9 wolbachia p
875	906	43.8	1	ADA22_HUMAN	Q9tp0k1 homo sapien	948	35	43.2	196	2	Q5T035_HUMAN	Q5t035 homo sapien
876	929	43.8	2	Q5PQW1_RAT	Q5pqw1 rattus norv	949	35	43.2	196	2	O30587_HELPY	O30587 helicobacte
877	999	43.8	2	Q4QHU9_LEIMA	Q4qhu9 leishmania	950	35	43.2	198	2	Q9NNV9_PLAFA	Q9nnv9 plasmodium
878	1008	43.8	2	Q8IVQ8_HUMAN	Q8ivq8 homo sapien	951	35	43.2	204	2	Q50XN0_ENTHI	Q50xn0 entamoeba h
879	1067	43.8	2	Q17517_CAEEL	Q17517 caenorhabdi	952	35	43.2	204	2	O5NT18_ENTHI	O5ntl8 entamoeba h
880	1067	43.8	2	Q7JMF9_CAEEL	Q7jmf9 caenorhabdi	953	35	43.2	204	2	Q7XK33_ORYSA	Q7xk33 oryza sativ
881	1097	43.8	1	KPC1_CANAL	P43057 candida alb	954	35	43.2	207	2	Q5BVS9_SCHJA	Q5bvs9 schistosoma
882	1097	43.8	2	Q5ANK2_CANAL	Q5ank2 candida alb	955	35	43.2	208	2	Q5UC47_HELPY	Q5uc47 helicobacte
883	1138	43.8	2	Q61ZX3_CAEBR	Q61zx3 caenorhabdi	956	35	43.2	209	2	Q8RJF5_HELPY	Q8rjf5 helicobacte
884	1149	43.8	2	Q6ZPJ4_MOUSE	Q6zpj4 mus musculu	957	35	43.2	210	2	O30585_HELPY	O30585 helicobacte
885	1151	43.8	1	XPO4_HUMAN	Q9c0e2 homo sapien	958	35	43.2	210	2	Q5UC40_HELPY	Q5uc40 helicobacte
886	1151	43.8	1	XPO4_MOUSE	Q9esj0 mus musculu	959	35	43.2	210	2	Q5UC42_HELPY	Q5uc42 helicobacte
887	1151	43.8	2	Q5VUZ5_HUMAN	Q5vuz5 homo sapien	960	35	43.2	210	2	Q5UC43_HELPY	Q5uc43 helicobacte
888	1154	43.8	2	Q5ZMR9_CHICK	Q5zmr9 gallus gall	961	35	43.2	210	2	Q5UC44_HELPY	Q5uc44 helicobacte
889	1412	43.8	2	O64612_ARATH	O64612 arabidopsis	962	35	43.2	210	2	Q5UC48_HELPY	Q5uc48 helicobacte
890	1506	43.8	2	Q6P6B9_HUMAN	Q6p6b9 homo sapien	963	35	43.2	210	2	Q5UC49_HELPY	Q5uc49 helicobacte
891	1566	43.8	2	Q9P2R6_HUMAN	Q9p2r6 homo sapien	964	35	43.2	210	2	Q5UC50_HELPY	Q5uc50 helicobacte
892	1566	43.8	2	Q5VXL9_HUMAN	Q5vxl9 homo sapien	965	35	43.2	210	2	Q5V9R5_HELPY	Q5v9r5 helicobacte
893	1680	43.8	2	Q4WS60_ASPFU	Q4ws60 aspergillus	966	35	43.2	210	2	Q5V9R6_HELPY	Q5v9r6 helicobacte
894	1692	43.8	2	Q6MYA4_ASPFU	Q6mya4 aspergillus	967	35	43.2	210	2	Q8RJ56_HELPY	Q8rj56 helicobacte
895	2017	43.8	2	Q582X4_9TRYP	Q582x4 trypanosoma	968	35	43.2	210	2	Q8RJ83_HELPY	Q8rj83 helicobacte
896	4151	43.8	2	Q57UX4_9TRYP	Q57ux4 trypanosoma	969	35	43.2	210	2	Q8RJE5_HELPY	Q8rje5 helicobacte
897	9579	43.8	2	Q4HWW4_GIBZE	Q4hww4 gibberella	970	35	43.2	210	2	Q8RSS2_HELPY	Q8rss2 helicobacte
898	10495	43.8	2	Q4RE92_TETNG	Q4hre92 tetraodon n	971	35	43.2	210	2	Q8VTN5_HELPY	Q8vtn5 helicobacte
899	69	43.2	2	Q91TP7_TUHV1	Q91tp7 tupaiid her	972	35	43.2	210	2	Q8VTN6_HELPY	Q8vtn6 helicobacte
900	83	43.2	2	Q9PFN0_XYLFA	Q9pfno xylella fas	973	35	43.2	210	2	Q8VTN7_HELPY	Q8vtn7 helicobacte
901	101	43.2	2	Q5V0H5_HALMA	Q5v0h5 haloarcula	974	35	43.2	210	2	Q8VTN9_HELPY	Q8vtn9 helicobacte
902	103	43.2	2	Q71Z55_LISMF	Q71z55 listeria mo	975	35	43.2	210	2	Q8VTP1_HELPY	Q8vtp1 helicobacte
903	103	43.2	2	Q8Y6S1_LISMO	Q8y6s1 listeria mo	976	35	43.2	210	2	Q8VTP2_HELPY	Q8vtp2 helicobacte
904	103	43.2	2	Q92B97_LISIN	Q92b97 listeria in	977	35	43.2	210	2	Q8VTP3_HELPY	Q8vtp3 helicobacte
905	103	43.2	2	Q8YUR8_ANASP	Q8yur8 anabaena sp	978	35	43.2	210	2	Q8VTP5_HELPY	Q8vtp5 helicobacte
906	106	43.2	2	Q4J0F7_AZOVI	Q4j0f7 azotobacter	979	35	43.2	210	2	Q8VTP6_HELPY	Q8vtp6 helicobacte
907	107	43.2	2	Q4K3V8_PSEF5	Q4k3v8 pseudomonas	980	35	43.2	210	2	Q8VTP7_HELPY	Q8vtp7 helicobacte

981 35 43.2 210 2 Q9FBD1_HELPY Q9fbd1 helicobacte
982 35 43.2 210 2 Q9FBD2_HELPY Q9fbd2 helicobacte
983 35 43.2 210 2 Q9FBD3_HELPY Q9fbd3 helicobacte
984 35 43.2 210 2 Q9FBD4_HELPY Q9fbd4 helicobacte
985 35 43.2 210 2 Q9FBD5_HELPY Q9fbd5 helicobacte
986 35 43.2 210 2 Q9FBE0_HELPY Q9fbe0 helicobacte
987 35 43.2 210 2 Q9FBE1_HELPY Q9fbe1 helicobacte
988 35 43.2 210 2 Q9FBE3_HELPY Q9fbe3 helicobacte
989 35 43.2 210 2 Q9FBE4_HELPY Q9fbe4 helicobacte
990 35 43.2 210 2 Q9FBE5_HELPY Q9fbe5 helicobacte
991 35 43.2 210 2 Q9FBE7_HELPY Q9fbe7 helicobacte
992 35 43.2 210 2 Q9FBE9_HELPY Q9fbe9 helicobacte
993 35 43.2 210 2 Q9FBF2_HELPY Q9fbf2 helicobacte
994 35 43.2 210 2 Q9FBD8_HELPY Q9fbd8 helicobacte
995 35 43.2 210 2 Q4TTP6_HELPY Q4ttp6 helicobacte
996 35 43.2 210 2 Q4TTP7_HELPY Q4ttp7 helicobacte
997 35 43.2 210 2 Q4TTP8_HELPY Q4ttp8 helicobacte
998 35 43.2 210 2 Q4TTP9_HELPY Q4ttp9 helicobacte
999 35 43.2 210 2 Q4TTQ3_HELPY Q4ttq3 helicobacte
1000 35 43.2 210 2 Q4TTQ4_HELPY Q4ttq4 helicobacte

ALIGNMENTS

RESULT 1
Q6CNV5_KLUULA
ID Q6CNV5_KLUULA PRELIMINARY; PRT; 488 AA.
AC Q6CNV5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|Q04673 Saccharomyces cerevisiae YLR005w SSL1 TFIIH
DE subunit.
GN OrderedLocusNames=KLLA0E09592g;
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
DR EMBL; CR382125; CAG99471.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR004595; Ssl1.
DR InterPro; IPR007198; Ssl1 like.
DR InterPro; IPR012170; TFIIH_SSL1.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04056; Ssl1; 1.
DR PIRSF; PIRSF015919; TFIIH_SSL1; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMs; TIGR00622; ssl1; 1.
DR PROSITE; PS50234; VWFA; 1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 488 AA; 54791 MW; E3BD0DC4C55FF08D CRC64;

Query Match 60.5%; Score 49; DB 2; Length 488;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
Db 360 CACHSTLVHGGYVC 373
|:|:|:|:|:|:|

RESULT 2
Q5S4N1_HUMAN
ID Q5S4N1_HUMAN PRELIMINARY; PRT; 362 AA.
AC Q5S4N1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shi Z., Wang H., Feng E., Su G., Huang L.;
RT "New sequences related to infection with Shigella flexneri 2a.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AY776161; AAV52794.1; -; mRNA.
DR Ensembl; ENSG00000124487; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;
KW Nucleotide-binding.
FT NON_TER 1 1
FT NON_TER 362 362
SQ SEQUENCE 362 AA; 40187 MW; 3EB871DFC708980B CRC64;

Query Match 59.3%; Score 48; DB 2; Length 362;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
Db 173 SLEDFLYHEGYAC 185
|:|:|:|:|:|:|

RESULT 3
Q9FFN6_ARATH
ID Q9FFN6_ARATH PRELIMINARY; PRT; 365 AA.
AC Q9FFN6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gb|AAD32776.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT p1 clones.";
RL DNA Res. 4:215-230(1997).
DR EMBL; AB005234; BAB10471.1; -; Genomic DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 365 AA; 39838 MW; 0E449FDD87533A02 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 365;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGYYVC 14
|:|: ||:| | |
Db 152 CALKWFLNHGSTVC 165

RESULT 4
Q8GUG6 ARATH
ID Q8GUG6_ARATH PRELIMINARY; PRT; 367 AA.
AC Q8GUG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein At5g63780.
GN Name=At5g63780;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinn P., Tang C.C., Toroumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J., Theologis A.,
RA Davis R.W.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT002521; AAC00881.1; -; mRNA.
DR EMBL; BT008456; AAP37815.1; -; mRNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.

KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ SEQUENCE 367 AA; 40111 MW; 1C846DFEB4BIACE7 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 367;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGYYVC 14
|:|: ||:| | |
Db 152 CALKWFLNHGSTVC 165

RESULT 5
Q4R9A4 MACFA
ID Q4R9A4_MACFA PRELIMINARY; PRT; 480 AA.
AC Q4R9A4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: QtsA-10416, similar to human DEAD (Asp-Glu-Ala-Asp)
DE box polypeptide 3, X-linked(DDX3X), transcript variant 1,..
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AB168192; BAE00317.1; -; mRNA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
SQ SEQUENCE 480 AA; 52923 MW; CD5F3F1F7557A809 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 480;
Best Local Similarity 61.5%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYYVC 14
||:| | | | |
Db 289 SLEDFLYHEGYAC 301

RESULT 6
DDX3Y PONPY
ID DDX3Y_PONPY STANDARD; PRT; 658 AA.
AC Q5RF43;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN Name=DDX3Y;
OS Pongo pygmaeus (Orangutan).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]_NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Kidney;
RC The German cDNA consortium;
RG Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: Probable ATP-dependent RNA helicase. May play a role in
CC spermatogenesis (By similarity).
CC -!- SUBCELLULAR LOCATION: Shuttles between the nucleus and the
CC cytoplasm in an XPO1-dependent manner (By similarity).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; CR857318; CAH89614.1; -; mRNA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; DNA-binding; Helicase; Hydrolase; Nuclear protein;
KW Nucleotide-binding; RNA-binding.
FT NP BIND 222 229 ATP (Potential).
FT MOTIF 345 348 DEAD box.
SQ SEQUENCE 658 AA; 73043 MW; 21237B600A5FDD6D CRC64;

Query Match 59.3%; Score 48; DB 1; Length 658;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 SLQEFSLSHGGYVC 14
DB 454 SLEDFLYHEGYAC 466

RESULT 7
DDX3Y HUMAN STANDARD; PRT; 660 AA.
AC O15523; Q8IYV7;
DT 15-JUL-1999 (Rel. 38, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN Name=DDX3Y; Synonyms=DBY;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_NUCLEOTIDE SEQUENCE [MRNA].
RP MEDLINE=98022381; PubMed=9381176; DOI=10.1126/science.278.5338.675;
RX Lahn B.T., Page D.C.;
RT "Functional coherence of the human Y chromosome.";
RL Science 278:675-680(1997).
RN [2]_NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Testis;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SUBCELLULAR LOCATION.
RX PubMed=15383328; DOI=10.1016/j.yexcr.2004.07.005;
RA Sekiguchi T., Iida H., Fukumura J., Nishimoto T.;
RT "Human DDX3Y, the Y-encoded isoform of RNA helicase DDX3, rescues a
RT hamster temperature-sensitive ET24 mutant cell line with a DDX3X
RT mutation.";
RL Exp. Cell Res. 300:213-222(2004).
RN [4]
RP TISSUE SPECIFICITY.
RX PubMed=15294876; DOI=10.1093/hmg/ddh240;
RA Ditton H.J., Zimmer J., Kamp C., Rajpert-De Meyts E., Vogt P.H.;
RT "The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is
RT limited to the male germ cells by translation control.";
RL Hum. Mol. Genet. 13:2333-2341(2004).
CC -!- FUNCTION: Probable ATP-dependent RNA helicase. May play a role in
CC spermatogenesis.
CC -!- SUBCELLULAR LOCATION: Shuttles between the nucleus and the
CC cytoplasm in an XPO1-dependent manner.
CC -!- TISSUE SPECIFICITY: Testis-specific. Expressed predominantly in
CC spermatogonia.
CC -!- DISEASE: DDX3Y is deleted in severe nonobstructive
CC hypospermatogenesis [MIM:400042].
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF000985; AAC51832.1; -; mRNA.
DR EMBL; AF000984; AAC51831.1; -; mRNA.
DR EMBL; BC034942; AAH34942.1; -; mRNA.
DR HSSP; P10081; 1FUK.
DR Ensembl; ENSG00000067048; Homo sapiens.
DR HGNC; HGNC:2699; DDX3Y.
DR MIM; 400010; -.
DR MIM; 400042; -.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; DNA-binding; Helicase; Hydrolase; Nuclear protein;
KW Nucleotide-binding; RNA-binding.
FT NP BIND 222 229 ATP (Potential).
FT MOTIF 345 348 DEAD box.
FT CONFLICT 46 46 R -> K (in Ref. 1).
FT CONFLICT 219 219 M -> V (in Ref. 1).
FT CONFLICT 628 628 N -> D (in Ref. 1).
SQ SEQUENCE 660 AA; 73154 MW; 0C370E9367952AB6 CRC64;

Query Match 59.3%; Score 48; DB 1; Length 660;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 SLQEFSLSHGGYVC 14

Db 454 SLEDFLYHEGYAC 466

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus. NCBI_TaxID=10090;

[1] NUCLEOTIDE SEQUENCE. TISSUE=Testis; MEDLINE=89249320; PubMed=2720782; DOI=10.1016/0092-8674(89)90125-6; Leroy P., Alzari P., Sassoon D., Wolgemuth D., Fellous M.; "The protein encoded by a murine male germ cell-specific transcript is a putative ATP-dependent RNA helicase."; Cell 57:549-559(1989).

[2] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. STRAIN=C57BL/6J; TISSUE=Testis; MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266; Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Corbani L.E., Cousins S., Blake J.A., Bradt D., Brusic V., Chothia C., Forrest A., Frazer K.S., Dalla E., Dragani T.A., Fletcher C.F., Godzik A., Gough J., Gaasterland T., Gariboldi M., Gissi C., Jackson I.J., Jarvis E.D., Grimmond S., Gustincich S., Hirokawa N., Kedzierski R.M., King B.L., Kanai A., Kawaji H., Kawasawa Y., Lee Y., Lenhard B., Lyons P.A., Konagaya A., Kurochkin I.V., Marchionni L., McKenzie L., Miki H., Maglott D.R., Maltais L., Okido T., Pavan W.J., Pertea G., Pesole G., Nagashima T., Numata K., Pontius J.U., Qi D., Ramachandran S., Petrovsky N., Pillai R., Reed J.C., Reid J., Ring B.Z., Ringwald M., Ravasi T., Reed J.C., Schneider C., Semple C.A., Setou M., Shimada K., Sandelin A., Schmeider C., Taylor M.S., Teasdale R.D., Tomita M., Sultana R., Takenaka Y., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Verardo R., Wagner L., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Wilming L.G., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Yuan Z., Zavolan M., Konno H., Nakamura M., Sakazume N., Sato K., Hirozane-Kishikawa T., Kawai J., Aizawa K., Arakawa T., Fukuda S., Shiraki T., Waki K., Imotani K., Ishii Y., Itoh M., Kagawa I., Hara A., Hashizume W., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Miyazaki A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.; "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; Nature 420:563-573(2002).

CC -!- FUNCTION: Putative ATP-dependent RNA helicase. Possible role in a key step of the spermatogenic process. CC -!- TISSUE SPECIFICITY: Testis. CC -!- DEVELOPMENTAL STAGE: High levels of PL10 during the meiotic and haploid stages of spermatogenesis. CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3 subfamily.

----- This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; J04847; AAA39942.1; -; mRNA. DR EMBL; AK029542; BAC26505.1; -; mRNA. DR PIR; A32378; A32378. DR HSSP; Q58083; 1HV8. DR Ensembl; ENSMUSG00000039224; Mus musculus. DR MGI; MGI:91842; D1Pas1. DR InterPro; IPR001410; DEAD. DR InterPro; IPR011545; DEAD/DEAH_N. DR InterPro; IPR000629; DEAD box. DR InterPro; IPR001650; Helicase_C. DR Pfam; PF00270; DEAD; 1. DR Pfam; PF00271; Helicase C; 1. DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1. DR ATP-binding; Developmental protein; Differentiation; DNA-binding;

Db 454 SLEDFLYHEGYAC 466

RESULT 8 DDX3Y_PANTR STANDARD; PRT; 660 AA. AC Q6GVM6; 25-OCT-2004 (Rel. 45, Created) DT 25-OCT-2004 (Rel. 45, Last sequence update) DT 13-SEP-2005 (Rel. 48, Last annotation update) DE DEAD-box protein 3, Y-chromosomal. GN Name=DDX3Y; OS Pan troglodytes (Chimpanzee).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; OC Pan. NCBI_TaxID=9598; [1] NUCLEOTIDE SEQUENCE. TISSUE=Testis; Hughes J.F., Pyntikova T., Skaletsky H., Minx P.J., Rozen S., Wilson R.K., Page D.C.; "The DNA sequence of the chimpanzee Y chromosome."; Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases. CC -!- FUNCTION: Probable ATP-dependent RNA helicase. May play a role in spermatogenesis (By similarity). CC -!- SUBCELLULAR LOCATION: Shuttles between the nucleus and the cytoplasm in an XPO1-dependent manner (By similarity). CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3 subfamily.

----- This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AY633112; AAT46349.1; -; mRNA. DR InterPro; IPR001410; DEAD. DR InterPro; IPR011545; DEAD/DEAH_N. DR InterPro; IPR000629; DEAD box. DR InterPro; IPR001650; Helicase_C. DR Pfam; PF00270; DEAD; 1. DR Pfam; PF00271; Helicase C; 1. DR SMART; SM00487; DEXDC; 1. DR SMART; SM00490; HELICC; 1. DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1. KW ATP-binding; DNA-binding; Helicase; Hydrolase; Nuclear protein; Nucleotide-binding; RNA-binding. FT NP_BIND 222 229 ATP (Potential). FT MOTIF 345 348 DEAD box. FT SEQUENCE 660 AA; 73191 MW; B84FDEC14235AC74 CRC64;

Query Match 59.3%; Score 48; DB 1; Length 660; Best Local Similarity 61.5%; Pred. No. 22; Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 2 SLOEFLSHGGYVC 14 Db 454 SLEDFLYHEGYAC 466

RESULT 9 PL10_MOUSE STANDARD; PRT; 660 AA. ID PL10_MOUSE AC P16381; 01-AUG-1990 (Rel. 15, Created) DT 01-AUG-1990 (Rel. 15, Last sequence update) DT 13-SEP-2005 (Rel. 48, Last annotation update) DE Putative ATP-dependent RNA helicase PL10. GN Name=D1Pas1; Synonyms=Pl10; OS Mus musculus (Mouse).

KW Helicase; Hydrolase; Nucleotide-binding; RNA-binding; Spermatogenesis.
FT NP_BIND 223 230 ATP (Potential).
FT DNA_BIND 495 514 Potential.
FT MOTIF 346 349 DEAD box.
FT COMBIAS 581 660 Gly/Ser-rich.
SQ SEQUENCE 660 AA; 73141 MW; 50AD4E6A131AE603 CRC64;

Query Match 59.3%; Score 48; DB 1; Length 660;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
||:|||||
Db 455 SLEDFLYHEGYAC 467

RESULT 10
DDX3X_HUMAN STANDARD; PRT; 661 AA.
ID DDX3X_HUMAN
AC O00571; O15536;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, X-chromosomal (Helicase-like protein 2) (HLP2)
DE (DEAD-box, X isoform).
GN Name=DDX3X; Synonyms=DBX, DDX3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hippocampus, and Liver;
RA Chung J., Lee S.-G., Song K.;
RT "Identification of a human homolog of a putative RNA helicase gene
RT (mDEAD3) expressed in mouse erythroid cells.";
RL Korean J. Biochem. 27:193-197(1995).

RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, SUBCELLULAR LOCATION, AND
RP INTERACTION WITH HCV CORE PROTEIN.
RP TISSUE=Liver;
RX MEDLINE=99263161; PubMed=10329544; DOI=10.1006/viro.1999.9659;
RA Owsianka A.M., Patel A.H.;
RT "Hepatitis C virus core protein interacts with a human DEAD box
RT protein DDX3.";
RL Virology 257:330-340(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98022381; PubMed=9381176; DOI=10.1126/science.278.5338.675;
RA Lahn B.T., Page D.C.;
RT "Functional coherence of the human Y chromosome.";
RL Science 278:675-680(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP PROTEIN SEQUENCE OF 1-9, AND ACETYLATION.
RX MEDLINE=20318637; PubMed=10859333; DOI=10.1084/jem.191.12.2083;
RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA Lopez de Castro J.A.;
RT "An N-acetylated natural ligand of human histocompatibility leukocyte
RT antigen (HLA)-B39. Classical major histocompatibility complex class I
RT proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL J. Exp. Med. 191:2083-2092(2000).
RN [6]
RP FUNCTION IN HIV-1 RNA EXPORT AND REPLICATION, IDENTIFICATION IN A
RP COMPLEX WITH XPO1 AND REV, INTERACTION WITH XPO1, MUTAGENESIS OF
RP LYS-229 AND SER-381, AND SUBCELLULAR LOCATION.
RX PubMed=15507209; DOI=10.1016/j.cell.2004.09.029;
RA Yedavalli V.S., Neuveut C., Chi Y.-H., Kleiman L., Jeang K.-T.;
RT "Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export
RT function.";
RL Cell 119:381-392(2004).
CC -!- FUNCTION: ATP-dependent RNA helicase. Acts as a cofactor for XPO1-
CC mediated nuclear export of incompletely spliced HIV-1 Rev RNAs.
CC Also involved in HIV-1 replication. Interacts specifically with
CC hepatitis C virus core protein resulting in a change in
CC intracellular location.
CC -!- SUBUNIT: Found in a complex with Rev and XPO1. Interacts with
CC XPO1. Interacts with HCV core protein.
CC -!- INTERACTION:
CC Q00653:NFKB2; NbExp=1; IntAct=EBI-353779, EBI-307326;
CC -!- SUBCELLULAR LOCATION: Located predominantly in nuclear speckles
CC and, at low levels, throughout the cytoplasm. Located to the outer
CC side of nuclear pore complexes (NPC). Shuttles between the nucleus
CC and the cytoplasm in a XPO1-dependent manner.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U50553; AAB95637.1; -; mRNA.
DR EMBL; AF061337; AAC34298.1; -; mRNA.
DR EMBL; AF000983; AAC51830.1; -; mRNA.
DR EMBL; AF000982; AAC51829.1; -; mRNA.
DR EMBL; BC011819; AAH11819.1; -; mRNA.
DR HSSP; Q58083; 1HV8.
DR IntAct; O00571; -.
DR SWISS-2DPAGE; O00571; HUMAN.
DR Ensembl; ENSG00000124487; Homo sapiens.
DR HGNC; HGNC:2745; DDX3X.
DR H-InvDB; HIX0016737; -.
DR MIM; 300160; -.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0004004; F:ATP-dependent RNA helicase activity; TAS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Acetylation; ATP-binding; Direct protein sequencing; DNA-binding;
KW Helicase; Hydrolase; Nuclear protein; Nucleotide-binding;
KW Phosphorylation; RNA-binding.
FT INIT MET 0
FT NP_BIND 223 230 ATP (Potential).
FT

FT REGION 259 516 Necessary for interaction with XPO1.
FT MOTIF 346 349 DEAD box.
FT COMPBIAS 581 661 Gly/Ser-rich.
FT MOD_RES 1 1 N-acetylserine.
FT MUTAGEN 73 73 Phosphoserine (By similarity).
FT 229 K->E: Abolishes ATPase activity and RNA-unwinding activity.
FT MUTAGEN 381 381 S->L: Abolishes ATPase activity and RNA-unwinding activity.
FT CONFLICT 49 49 K -> R (in Ref. 3).
SQ SEQUENCE 661 AA; 73112 MW; F0F03DD8FBC00A65 CRC64;
Query Match 59.3%; Score 48; DB 1; Length 661;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLQEFSLSHGGYVC 14
Db 455 SLEDFLYHEGYAC 467
RESULT 11
DDX3X_MOUSE STANDARD; PRT; 661 AA.
ID DDX3X_MOUSE 009060; O09143;
AC Q62167; O09060; O09143;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, X-chromosomal (DEAD-box RNA helicase DEAD3)
DE (mDEAD3) (Embryonic RNA helicase) (DlPas1-related sequence 2).
GN Name=Ddx3x; Synonyms=DlPas1-rs2, Ddx3, Dead3, Erh;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6, and DBA;
RX MEDLINE=97104282; PubMed=8948440;
RA Sowden J.C., Putt W., Morrison K., Beddington R., Edwards Y.;
RT "The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of RNA helicases.";
RL Biochem. J. 308:839-846(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Erythroleukemia;
RX MEDLINE=94192995; PubMed=8144024; DOI=10.1016/0378-1119(94)90541-X;
RA Gee S.L., Conboy J.G.;
RT "Mouse erythroid cells express multiple putative RNA helicase genes exhibiting high sequence conservation from yeast to mammals.";
RL Gene 140:171-177(1994).
RN [3]
RP PROTEIN SEQUENCE OF 1-9, AND ACETYLATION.
RX MEDLINE=20318637; PubMed=10859333; DOI=10.1084/jem.191.12.2083;
RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA Lopez de Castro J.A.;
RT "An N-acetylated natural ligand of human histocompatibility leukocyte antigen (HLA)-B39. Classical major histocompatibility complex class I proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL J. Exp. Med. 191:2083-2092(2000).
RN [4]
RP PHOSPHORYLATION SITE SER-73.
RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
RA Shu H., Chen S., Bi Q., Mumby M., Brekken D.L.;
RT "Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line.";
RL Mol. Cell. Proteomics 3:279-286(2004).
CC -!- FUNCTION: Putative ATP-dependent RNA helicase. It may play a role in translational activation of mRNA in the oocyte and early embryo.
CC -!- SUBUNIT: Interacts with XPO1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Located predominantly in nuclear speckles

CC and, at low levels, throughout the cytoplasm. Located to the outer side of nuclear pore complexes (NPC). Shuttles between the nucleus and the cytoplasm in a XPO1-dependent manner (By similarity).
CC -!- TISSUE SPECIFICITY: Developmentally regulated.
CC -!- DEVELOPMENTAL STAGE: Expressed in oocytes. Ubiquitously found in 9 days post-conception embryo, at later stages it is restricted to brain and kidney.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3 subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL; Z38117; CAA86261.1; -; mRNA.
DR EMBL; L25126; AAA53630.1; -; mRNA.
DR PIR; I84741; I84741.
DR HSSP; Q58083; LHV8.
DR Ensembl; ENSMUSG00000000787; Mus musculus.
DR MGI; MGI:103064; Ddx3x.
DR GO; GO:0005737; C:cytoplasm; ISS.
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0004004; F:ATP-dependent RNA helicase activity; ISS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Acetylation; ATP-binding; Direct protein sequencing; DNA-binding; Helicase; Hydrolase; Nuclear protein; Nucleotide-binding; Phosphorylation; RNA-binding.
FT INIT_MET 0 0
FT NP_BIND 223 230 ATP (Potential).
FT REGION 259 516 Necessary for interaction with XPO1 (By similarity).
FT MOTIF 346 349 DEAD box.
FT COMPBIAS 581 661 Gly/Ser-rich.
FT COMPBIAS 608 615 Poly-Ser.
FT COMPBIAS 623 629 Poly-Gly.
FT COMPBIAS 632 640 Poly-Gly.
FT MOD_RES 1 1 N-acetylserine.
FT MOD_RES 73 73 Phosphoserine.
SQ SEQUENCE 661 AA; 72970 MW; A1E1FAAB5D19F57B CRC64;
Query Match 59.3%; Score 48; DB 1; Length 661;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLQEFSLSHGGYVC 14
Db 455 SLEDFLYHEGYAC 467
RESULT 12
Q5JSI3_HUMAN
ID Q5JSI3_HUMAN PRELIMINARY; PRT; 662 AA.
AC Q5JSI3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked.
GN Name=DDX3X; ORFNames=RP1-169I5.3-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.


```
RT biogeographic implications of an analysis of nuclear DNA sequence
RT data.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AY055971; AAL18088.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
FT NON_TER 202
FT NON_TER 202
SQ SEQUENCE 202 AA; 21898 MW; BFC6D57F86680A9D CRC64;

Query Match 56.8%; Score 46; DB 2; Length 202;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 CS-LQEFLSHGGYVC 14
||| | | | | | | |
DB 164 CSQKLEEGLSHSSYVC 179

RESULT 16
Q7XK97 ORYSA
ID Q7XK97_ORYSA PRELIMINARY; PRT; 358 AA.
AC Q7XK97;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNBb0020J19.13 protein.
GN Name=OSJNBb0020J19.13;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606656; CAE05784.2; -; Genomic_DNA.
DR Gramene; Q7XK97; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001309; ICE_p20.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PROSITE; PS01121; CASPASE_HIS; UNKNOWN 1.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 358 AA; 38844 MW; F89A0D01AC61771D CRC64;
```

```
Query Match 56.8%; Score 46; DB 2; Length 358;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
| | | | | | | |
DB 120 CALKWFITHGSTVC 133

RESULT 17
Q69BL0.MANSE
ID Q69BL0.MANSE PRELIMINARY; PRT; 666 AA.
AC Q69BL0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Pattern recognition serine proteinase precursor.
GN Name=PRSP;
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingoidea;
OC Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15190055; DOI=10.1074/jbc.M404584200;
RA Ji C., Wang Y., Guo X., Hartson S., Jiang H.;
RT "A pattern recognition serine proteinase triggers the prophenoloxidase
RT activation cascade in the tobacco hornworm, Manduca sexta.";
RL J. Biol. Chem. 279:34101-34106(2004).
DR EMBL; AY380790; AAR29602.1; -; mRNA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00057; Ldl_recept_a; 5.
DR Pfam; PF00084; Sushi; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00192; LDLa; 5.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 4.
DR PROSITE; PS50068; LDLRA_2; 3.
DR PROSITE; PS50923; SUSHI; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 666 pattern recognition serine proteinase.
SQ SEQUENCE 666 AA; 73670 MW; 880439BC41B9238F CRC64;

Query Match 56.8%; Score 46; DB 2; Length 666;
Best Local Similarity 61.5%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYV 13
| | | | | | | |
DB 258 CVLPPEYEHGGYV 270

RESULT 18
Q93ZD2 ARATH
ID Q93ZD2_ARATH PRELIMINARY; PRT; 363 AA.
AC Q93ZD2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```


DE AT5g63780/MBK5_26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057621; AAL1416.1; -; mRNA.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR011016; RINGV.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00744; RINGv; 1.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 363 AA; 39458 MW; 8BA2C124FB76BE8E CRC64;

Query Match 55.6%; Score 45; DB 2; Length 363;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGYYVC 14
DB 140 CALKWFVNHGSTVC 153

RESULT 19
Q9CSA2 ARATH PRELIMINARY; PRT; 363 AA.
AC Q9CSA2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein At5g08750 (Hypothetical protein
DE At5g08750/T2K12_100).
GN Name=At5g08750; Synonyms=At5g08750/T2K12_100;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590346; CAC35880.1; -; Genomic_DNA.
DR EMBL; AK117992; BAC42627.1; -; mRNA.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR011016; RINGV.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00744; RINGv; 1.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ SEQUENCE 363 AA; 39430 MW; 6B1A8E05A45F7281 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 363;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGYYVC 14
DB 140 CALKWFVNHGSTVC 153

RESULT 20
Q5ER66 MONAL PRELIMINARY; PRT; 376 AA.
ID Q5ER66 MONAL PRELIMINARY;
AC Q5ER66;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DEAD box RNA helicase-PL10A (Fragment).
OS Monopterus albus (Swamp eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Synbranchiformes;
OC Synbranchidae; Monopterus.
OX NCBI_TaxID=43700;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Maoyu P., Ping S.;
RT "Cloning and Sequence Analysis of Monopterus albus DEAD-box Family
RT Gene PL10.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AY786203; AAW78518.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding.
FT NON_TER 1
FT NON_TER 376
SQ SEQUENCE 376 AA; 42248 MW; B42BA83E0B4BF859 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 376;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYYVC 14
DB 304 ALEDFLYHEGYAC 316

RESULT 21
Q518T7 ENTHI PRELIMINARY; PRT; 448 AA.
ID Q518T7 ENTHI PRELIMINARY;
AC Q518T7;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 5'-3' exonuclease, putative.
GN ORFNames=45.t00020;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000190; EAL49366.1; -; Genomic_DNA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR012282; Cytochrome c R.
DR InterPro; IPR012351; Cytokine_4_hlx.
DR InterPro; IPR004859; Put_53exo.
DR InterPro; IPR012335; Thioredoxin_fold.
DR Pfam; PF03159; XRN_N; 1.
KW Exonuclease.
SQ SEQUENCE 448 AA; 52207 MW; C860798270121D04 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 448;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QEFLSHGQYV 13
:|||||:
Db 345 REFLPHGGYI 354

RESULT 22
RS1_CHLMU STANDARD; PRT; 570 AA.
ID RS1_CHLMU
AC P38016;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 30S ribosomal protein S1 (70 kDa antigen).
GN Name=rpSA; OrderedLocusNames=TC0373;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-180.

RC STRAIN=MoPn;
RX MEDLINE=89123039; PubMed=2644193;
RA Sardinia L.M., Engel J.N., Ganem D.;
RT "Chlamydial gene encoding a 70-kilodalton antigen in Escherichia coli:
RT analysis of expression signals and identification of the gene
RT product.";
RL J. Bacteriol. 171:335-341(1989).
CC -!- FUNCTION: Binds mRNA; thus facilitating recognition of the
CC initiation point. It is needed to translate mRNA with a short
CC Shine-Dalgarno (SD) purine-rich sequence (By similarity).
CC -!- SIMILARITY: Belongs to the ribosomal protein S1P family.
CC -!- SIMILARITY: Contains 6 S1 motif domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE002304; AAF39231.1; ALT_INIT; Genomic_DNA.
CC EMBL; M23000; AAA23167.1; -; Genomic_DNA.
DR PIR; A32246; A32246.
DR PIR; A81710; A81710.
DR HSSP; P05055; 1SRO.
DR Siena-2DPAGE; P38016; -.
DR TIGR; TC0373; -.
DR InterPro; IPR012340; OB_NA_bd_sub.
DR InterPro; IPR000110; Ribosomal_S1.
DR InterPro; IPR003029; S1_RNA_bd.
DR Pfam; PF00575; S1; 6.
DR PRINTS; PR00681; RIBOSOMALS1.
DR SMART; SM00316; S1; 6.
DR TIGRFAMs; TIGR00717; rpsA; 1.
DR PROSITE; PS50126; S1; 6.
KW Complete proteome; Repeat; Ribonucleoprotein; Ribosomal protein;
KW RNA-binding.
FT DOMAIN 52 116 S1 motif 1.
FT DOMAIN 134 199 S1 motif 2.
FT DOMAIN 220 288 S1 motif 3.
FT DOMAIN 305 375 S1 motif 4.
FT DOMAIN 392 462 S1 motif 5.
FT DOMAIN 479 548 S1 motif 6.
FT CONFLICT 139 139 G -> A (in Ref. 2).
FT CONFLICT 143 143 R -> P (in Ref. 2).
FT CONFLICT 177 180 GKVC -> RESL (in Ref. 2).
SQ SEQUENCE 570 AA; 63610 MW; 5715B2C711518A24 CRC64;

Query Match 55.6%; Score 45; DB 1; Length 570;
Best Local Similarity 63.6%; Pred. No. 62;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLQEFLSHGQY 12
|:|:|:|:|:
Db 546 SIKEFLAHGGH 556

RESULT 23
Q4KXN1_MOUSE
ID Q4KXN1_MOUSE PRELIMINARY; PRT; 611 AA.
AC Q4KXN1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 1100001A21Rik protein.
GN Name=1100001A21Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6Ncr; TISSUE=Hematopoietic Stem Cell;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altechul S.F., Zeebun B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6Ncr; TISSUE=Hematopoietic Stem Cell;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC098471; AAH98471.1; -; mRNA.
DR MGI; MGI:1913845; 1100001A21Rik.
SQ SEQUENCE 611 AA; 68927 MW; 21644C2E5CB270CD CRC64;

Query Match 55.6%; Score 45; DB 2; Length 611;
Best Local Similarity 58.3%; Pred. No. 67;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQEFLSHGQYVC 14
||:|:|:|:|
Db 243 LQDFQYGSYVC 254

RESULT 24
Q5F491_CHICK PRELIMINARY; PRT; 651 AA.
AC Q5F491;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_2a4;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Kotter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AJ851409; CAH65043.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;
KW Nucleotide-binding.
SQ SEQUENCE 651 AA; 72049 MW; 7C4584A7AA396834 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 651;
Best Local Similarity 53.8%; Pred. No. 72;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGQYVC 14
:|:|:|:|:|
Db 448 ALEDFLYHEGYAC 460

RESULT 25
Q8BIR2_MOUSE PRELIMINARY; PRT; 672 AA.
AC Q8BIR2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
DE library, clone:5430408G21 product:similar to CDNA FLJ14900 FIS, CLONE
DE PLACE1005176, HIGHLY SIMILAR TO HOMO SAPIENS HYPOTHALAMUS PROTEIN
DE HT001 MRNA.
GN Name=1100001A21Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

RA NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[4]
RA NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RA NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Head;
RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK030661; BAC27067.1; -; mRNA.
DR Ensembl; ENSMUSG00000032567; Mus musculus.
DR MGI; MGI:1913845; 1100001A21Rik.
DR GO; GO:0003677; F:nucleosome binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR InterPro; IPR000513; Exo_N_I.
SQ SEQUENCE 672 AA; 76370 MW; 51F219D759443543 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 672;
Best Local Similarity 58.3%; Pred. No. 74;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQEFSLSHGGYVC 14
Db 304 LQDFQYGSYVC 315
||:|:|:|
|:|:|:|
Q52L23 XENLA PRELIMINARY; PRT; 695 AA.
Q52L23;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8335;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Oocytes;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Oocytes;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Bouffard G.G.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Oocytes;
RC Klein S., Gerhard D.S.;
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the DEAD box helicase family.
CC EMBL; BC094097; AAH94097.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;
KW Nucleotide-binding.
SQ SEQUENCE 695 AA; 77045 MW; DC5C235C3C06689A CRC64;

Query Match 55.6%; Score 45; DB 2; Length 695;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYVC 14
Db 495 ALEDFLYHEGYAC 507
:|:|:|:|:|:|
|:|:|:|:|:|:|

RESULT 27
AN3 XENLA STANDARD; PRT; 697 AA.
ID AN3 XENLA
AC P24346;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Putative ATP-dependent RNA helicase An3.
GN Name=AN3;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.

OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91141586; PubMed=1996140; DOI=10.1038/349717a0;
RA Gururajan R., Perry-O'Keefe H., Melton D.A., Weeks D.L.;
RT "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
RT RNA helicase.";
RL Nature 349:717-719(1991).
CC -!- FUNCTION: Putative ATP-dependent RNA helicase.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X57328; CAA40605.1; -; mRNA.
DR PIR; S13654; S13654.
DR HSSP; P10081; 1FUK.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
FT NP_BIND 265 272 ATP (By similarity).
FT MOTIF 388 391 DEAD box.
FT COMEBIAS 623 697 Gly/Ser-rich.
SQ SEQUENCE 697 AA; 77303 MW; F3DD23EB60B2E2EF CRC64;

Query Match 55.6%; Score 45; DB 1; Length 697;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
:|:|:|:|:|:|
Db 497 ALEDFLYHEGYAC 509

RESULT 28
Q7ZXJ0_XENLA PRELIMINARY; PRT; 697 AA.
AC Q7ZXJ0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P110-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; BC044972; AAH44972.1; -; mRNA.
DR HSSP; Q58083; 1HV8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding.
SQ SEQUENCE 697 AA; 77133 MW; C6ED241A0024A384 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 697;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGYVC 14
:|:|:|:|:|:|
Db 497 ALEDFLYHEGYAC 509

RESULT 29
Q6P4J3_XENTR PRELIMINARY; PRT; 699 AA.
AC Q6P4J3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC76021.
GN Name=MGC76021;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

RN NUCLEOTIDE SEQUENCE.

RC TISSUE=Embryo;

RA Klein S., Gerhard D.S.;

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Belongs to the DEAD box helicase family.

DR EMBL; BC063374; AAH63374.1; -; mRNA.

DR Ensembl; ENSXETG0000015474; Xenopus tropicalis.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011545; DEAD/DEAH_N.

DR InterPro; IPR000629; DEAD_box.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase_C; 1.

DR SMART; SM00487; DEXDc; 1.

DR SMART; SM00490; HELICC; 1.

DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.

KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;

KW Nucleotide-binding.

SQ SEQUENCE 699 AA; 77467 MW; 2AF0AB4977CB2299 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 699;

Best Local Similarity 53.8%; Pred. No. 77;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGGYVC 14

:||| |

Db 498 ALEDFLYHEGYAC 510

RESULT 30

Q51FR1_ENTHI

ID Q51FR1_ENTHI PRELIMINARY; PRT; 948 AA.

AC Q51FR1;

DT 13-SEP-2005 (TReMBLrel. 31, Created)

DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)

DE 5'-3' exonuclease, putative.

GN ORFNames=5.t00013;

OS Entamoeba histolytica HM-1:IMSS.

OC Eukaryota; Entamoebidae; Entamoeba.

OX NCBI_TaxID=294381;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HM-1:IMSS;

RX PubMed=15729342; DOI=10.1038/nature03291;

RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,

RA Amadeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,

RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,

RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,

RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,

RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,

RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,

RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000025; EAL51687.1; -; Genomic_DNA.
KW Exonuclease.
SQ SEQUENCE 948 AA; 111242 MW; D1BB9F8186A8034C CRC64;

Query Match 55.6%; Score 45; DB 2; Length 948;

Best Local Similarity 70.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QEFLSHGGYV 13

:||| |

Db 386 REFLPHGGYI 395

RESULT 31

Q9FJE2_ARATH

ID Q9FJE2_ARATH PRELIMINARY; PRT; 651 AA.

AC Q9FJE2;

DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Tail-like non-LTR retroelement protein-like.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=9087489; PubMed=9872454;

RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,

RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.

RT Sequence features of the regions of 1,013,767 bp covered by sixteen

RT physically assigned P1 and TAC clones.";

RL DNA Res. 5:297-308(1998).

DR EMBL; AB015475; BAB08362.1; -; Genomic_DNA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0005489; F:electron transporter activity; IEA.

DR GO; GO:0005506; F:iron ion binding; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR011424; Cl_3.

DR InterPro; IPR002219; DAG_PE-bind.

DR InterPro; IPR004146; DC1.

DR InterPro; IPR001965; Znf_PHD.

DR Pfam; PF03107; Cl_2; 2.

DR Pfam; PF07649; Cl_3; 3.

DR SMART; SM00109; Cl_2.

DR SMART; SM00249; PHD; 4.

SQ SEQUENCE 651 AA; 72730 MW; AAFDD67F0D8B869A CRC64;

Query Match 54.9%; Score 44.5; DB 2; Length 651;

Best Local Similarity 50.0%; Pred. No. 87;

Matches 7; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CSLQEFLSHGGYVC 14

|::|: |

Db 29 CGVEDFV-YGGYVC 41

RESULT 32

Q5DELO_SCHJA

ID Q5DELO_SCHJA PRELIMINARY; PRT; 266 AA.

KW Kinase.
SQ SEQUENCE 342 AA; 37312 MW; F24DEB3FEBEE6F3C CRC64;
Query Match 54.3%; Score 44; DB 2; Length 319;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LQEFLSHGGYV 13
| | | | | : | | | :
Db 176 LQEFVNHGGVI 186

RESULT 35
Q7XBW0 ORYSA
ID Q7XBW0_ORYSA PRELIMINARY; PRT; 342 AA.
AC Q7XBW0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative inositol 1,3,4-trisphosphate 5/6-kinase.
GN ORFNames=OSJNBa0027L23.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017122; AAP55148.1; -; Genomic_DNA.
DR Gramene; Q7XBW0; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR008656; Insl34_P3_kin.
DR Pfam; PF05770; Insl34_P3_kin; I.
KW Kinase.
SQ SEQUENCE 342 AA; 38072 MW; 3B99D2AB5E46FB93 CRC64;
Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LQEFLSHGGYV 13
| | | | | : | | | :
Db 185 LQEFVNHGGVI 195

RESULT 36
Q84Y01 MAIZE
ID Q84Y01_MAIZE PRELIMINARY; PRT; 342 AA.
AC Q84Y01;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol phosphate kinase.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22474773; PubMed=12586875; DOI=10.1104/pp.014258;
RA Shi J., Wang H., Wu Y., Hazebroek J., Meeley R.B., Ertl D.S.;
RT "The maize low-phytic acid mutant lpa2 is caused by mutation in an
inositol phosphate kinase gene."
RL Plant Physiol. 131:507-515 (2003).
DR EMBL; AY172635; AAO17299.1; -; mRNA.
DR GO; GO:0016301; F:kinase activity; IEA.

KW Kinase.
SQ SEQUENCE 342 AA; 37312 MW; F24DEB3FEBEE6F3C CRC64;
Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LQEFLSHGGYV 13
| | | | | : | | | :
Db 186 LQEFVNHGGVI 196

RESULT 37
Q8W3H7 ORYSA
ID Q8W3H7_ORYSA PRELIMINARY; PRT; 342 AA.
AC Q8W3H7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol 1,3,4-trisphosphate 5/6-kinase.
GN Name=OSJNBa0027L23.5;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018929; AAL67584.1; -; Genomic_DNA.
DR Gramene; Q8W3H7; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR008656; Insl34_P3_kin.
DR Pfam; PF05770; Insl34_P3_kin; I.
KW Kinase.
SQ SEQUENCE 342 AA; 38072 MW; 3B99D2AB5E46FB93 CRC64;
Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LQEFLSHGGYV 13
| | | | | : | | | :
Db 185 LQEFVNHGGVI 195

RESULT 38
Q92M44 RHIME
ID Q92M44_RHIME PRELIMINARY; PRT; 348 AA.
AC Q92M44;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PUTATIVE LOW SPECIFICITY L-THREONINE ALDOLASE PROTEIN
(EC 4.1.2.5).
GN OrderedLocusNames=R02812; ORFNames=SMC04029;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47391.1; -; Genomic_DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR001597; Beta_elim_lyase.
DR Pfam; PF01212; Beta_elim_lyase; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 348 AA; 37858 MW; 044516F6B5D5CA78 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 348;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
| | | | | : |
Db 90 CGAPEFFSHGARLC 103

RESULT 39
Q98E45_RHILO
ID Q98E45_RHILO PRELIMINARY; PRT; 388 AA.
AC Q98E45;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE M114416 protein.
GN OrderedLocusNames=m114416;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000012; BAB51075.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006076; Fad_oxred.
DR Pfam; PF01266; DAO; 1.
KW Complete proteome.
SQ SEQUENCE 388 AA; 41940 MW; 77B6F256B17490F5 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 388;
Best Local Similarity 72.7%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGY 12
: | | | | |
Db 353 ALAEFLVHGGY 363

RESULT 40
Q95K70_MACFA
ID Q95K70_MACFA PRELIMINARY; PRT; 393 AA.
AC Q95K70;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cerebellum cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AB066509; BAB62196.1; -; mRNA.
DR HSSP; P08603; 1HFH.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS50923; SUSHI; 3.
KW Hypothetical protein.
SQ SEQUENCE 393 AA; 43314 MW; F964759AF2C75CF9 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 393;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
| | | | | : | |
Db 146 CPLPPMVSHGDFVC 159

RESULT 41
Q6AJD7_DESPS
ID Q6AJD7_DESPS PRELIMINARY; PRT; 413 AA.
AC Q6AJD7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Related to penicillin-binding protein 4 [Precursor].
GN OrderedLocusNames=DP2814;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
RA Rabus R., Ruepp A., Frickey T., Rattai T., Partmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG37543.1; -; Genomic DNA.
DR GO; GO:0004185; F:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000667; Peptidase_S13.
DR Pfam; PF02113; Peptidase_S13; 1.
DR PRINTS; PR00922; DADACEPTASE3.
DR TIGRFAMs; TIGR00666; PBP4; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 46188 MW; B88E4A9149E52902 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 413;
Best Local Similarity 72.7%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGY 12


```
Db          34 SLQEDIDHGGY 44
||||| : |||||
RESULT 42
MYCS_RAT
ID MYCS_RAT STANDARD; PRT; 429 AA.
AC P23999;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE S-myc protein.
GN Name=Mycs;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90083224; PubMed=2594755;
RA Sugiyama A., Kume A., Nemoto K., Lee S.Y., Asami Y., Nemoto F.,
RA Nishimura S., Kuchino Y.;
RT "Isolation and characterization of s-myc, a member of the rat myc gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9144-9148(1989).
CC -!- FUNCTION: Has apoptosis-inducing activity.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; M29069; AAA41645.1; -; Genomic_DNA.
PIR; A36220; A36220.
HSSP; P25912; 1HLO.
TRANSFAC; T03550; -.
Ensembl; ENSRNOG0000003085; Rattus norvegicus.
RGD; 3133; Mycs.
InterPro; IPR001092; HLH_basic.
InterPro; IPR002418; TF_Myc.
InterPro; IPR012682; TF_Myc_N.
Pfam; PF00010; HLH; 1.
Pfam; PF01056; Myc N; 1.
PRINTS; PR00044; LEUZIPP MYC.
SMART; SM00353; HLH; 1.
PROSITE; PS50888; HLH; 1.
KW Apoptosis; DNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 360 399 Helix-loop-helix motif.
FT DOMAIN 398 419 Leucine-zipper (Potential).
FT DNA_BIND 347 359 Basic motif.
FT MOD_RES 36 36 Phosphotyrosine (by Tyr-kinases) (By
FT similarity).
SQ SEQUENCE 429 AA; 47002 MW; 1E79FF90747FD0FA CRC64;
Query Match 54.3%; Score 44; DB 1; Length 429;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 2 SLQEFSLSHGGYV 13
||||| : |||||
Db 227 SLEDFLNSGYV 238
RESULT 43
Q7N4F2 PHOLL
ID Q7N4F2 PHOLL PRELIMINARY; PRT; 430 AA.
AC Q7N4F2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein.
GN OrderedLocusNames=plu2387;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571867; CAE14715.1; -; Genomic_DNA.
DR Photolist; plu2387; -.
DR InterPro; IPR007067; Phage_sheath_1.
DR Pfam; PF04984; Phage_sheath_1; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 46718 MW; 45A307E1169B400F CRC64;
Query Match 54.3%; Score 44; DB 2; Length 430;
Best Local Similarity 53.8%; Pred. No. 68;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 SLQEFSLSHGGYVC 14
:||| : ||||| ||
Db 99 AVQHYFSHGGGVC 111
RESULT 44
Q5R8M2 PONPY
ID Q5R8M2 PONPY PRELIMINARY; PRT; 489 AA.
AC Q5R8M2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469E106.
GN Name=DKFZp469E106;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR859729; CAH91888.1; -; mRNA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 4.
DR SMART; SM00032; CCP; 4.
DR PROSITE; PS50923; SUSHI; 4.
KW Hypothetical protein.
SQ SEQUENCE 489 AA; 53697 MW; 5241CA08FFE29990 CRC64;
Query Match 54.3%; Score 44; DB 2; Length 489;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CSLQEFSLSHGGYVC 14
||| : ||||| ||
Db 242 CPLPPMVSHGDFVC 255
```


DR InterPro; IPR000110; Ribosomal_S1.
DR InterPro; IPR003029; S1.
DR Pfam; PF00575; S1; 1.
DR PRINTS; PR00681; RIBOSOMALS1.
DR SMART; SM00316; S1; 1.
DR PROSITE; PS50126; S1; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 123 AA; 13376 MW; 993084CF4DDA319C CRC64;

Query Match 53.1%; Score 43; DB 2; Length 123;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLQEFLSHGG 11
|:|||||
Db 100 SIKEFLAHGG 109

RESULT 49
Q4RU12 TETNG
ID Q4RU12 TETNG PRELIMINARY; PRT; 219 AA.
AC Q4RU12;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 12 SCAP14996, whole genome shotgun sequence.
GN ORFNames=GSTENG0028997001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014996; CAG08120.1; -; Genomic DNA.
SQ SEQUENCE 219 AA; 24161 MW; 7F62B3CBC968EC01 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 219;
Best Local Similarity 72.7%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGGY 12
|:|||||
Db 167 SAQEFLTHEGY 177

RESULT 50
Q759D9 ASHGO
ID Q759D9_ASHGO PRELIMINARY; PRT; 469 AA.

AC Q759D9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ADR338Cp.
GN Name=ADR338C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AB016817; AAS52258.1; -; Genomic_DNA.
DR AGD; ADR338C; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR004595; Ssl1.
DR InterPro; IPR007198; Ssl1 like.
DR InterPro; IPR012170; TFIIF_SSL1.
DR InterPro; IPR002035; VWF A.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04056; Ssl1; 1.
DR PIRSF; PIRSF015919; TFIIF_SSL1; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMS; TIGR00622; ssl1; 1.
DR PROSITE; PS50234; VWFA; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 469 AA; 52811 MW; 44638920D54DDF54 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 469;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGGYVC 14
|:|||||
Db 345 CTCHSRLVHGGYFC 358

Search completed: May 12, 2006, 10:52:31
Job time : 130.077 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 72.6496 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-62
Perfect score: 56
Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	3	AAB17006 TPO-mimet
2	56	100.0	10	5	ABB72892 TPO mimet
3	56	100.0	10	7	ADJ73043 TPO mimet
4	56	100.0	10	8	ADJ52678 CH1 delet
5	56	100.0	10	8	ADJ51639 CH1 delet
6	56	100.0	12	2	AAW36787 Thrombopo
7	56	100.0	12	2	AAW36781 Thrombopo
8	56	100.0	12	4	AAU26014 Human thr
9	56	100.0	13	2	AAW36792 Thrombopo
10	56	100.0	13	2	AAW36779 Thrombopo
11	56	100.0	13	4	AAU26008 Human thr
12	56	100.0	13	4	AAU26012 Human thr
13	56	100.0	13	4	AAU26015 Human thr
14	56	100.0	13	7	ADJ73003 TPO mimet
15	56	100.0	13	8	ADJ52638 CH1 delet
16	56	100.0	13	8	ADJ51599 CH1 delet
17	56	100.0	14	2	AAW09463 Thrombopo
18	56	100.0	14	2	AAW09468 Thrombopo
19	56	100.0	14	2	AAW33030 Thrombopo
20	56	100.0	14	2	AAW33034 Thrombopo
21	56	100.0	14	2	AAW36782 Thrombopo
22	56	100.0	14	2	AAW36774 Thrombopo
23	56	100.0	14	2	AAW36788 Thrombopo
24	56	100.0	14	2	ADI24843 AF 12505

25	56	100.0	14	3	AAAY96515 Thrombopo
26	56	100.0	14	3	AAB16962 TPO-mimet
27	56	100.0	14	3	AAB16968 TPO-mimet
28	56	100.0	14	4	AAU26009 Human thr
29	56	100.0	14	4	AAU26013 Human thr
30	56	100.0	14	4	AAU26006 Human thr
31	56	100.0	14	4	AAU25827 Human thr
32	56	100.0	14	4	AAU26010 Human thr
33	56	100.0	14	4	AAU26004 Human thr
34	56	100.0	14	5	ABB72854 TPO mimet
35	56	100.0	14	5	ABB72853 TPO mimet
36	56	100.0	14	5	ABP51669 Thrombopo
37	56	100.0	14	5	AAE18011 Human lig
38	56	100.0	14	6	ABG71747 TPO recep
39	56	100.0	14	7	ABR62907 Thrombopo
40	56	100.0	14	7	ADC33697 Erythro
41	56	100.0	14	7	ADJ73004 TPO mimet
42	56	100.0	14	7	ADN59652 Thrombopo
43	56	100.0	14	8	ADJ52639 CH1 delet
44	56	100.0	14	8	ADJ51600 CH1 delet
45	56	100.0	14	8	ADL27293 Amino aci
46	56	100.0	14	8	ADM72503 TPO mimet
47	56	100.0	14	8	ADM72483 TPO mimet
48	56	100.0	14	8	ADM72495 TPO mimet
49	56	100.0	14	8	ADM72487 TPO mimet
50	56	100.0	14	8	ADM72497 TPO mimet
51	56	100.0	14	8	ADQ16584 Agonist T
52	56	100.0	14	8	ADT92482 Thrombopo
53	56	100.0	14	9	ADU70210 Thrombopo
54	56	100.0	14	9	ADU70206 Thrombopo
55	56	100.0	14	9	ADU75982 Peptide-b
56	56	100.0	14	9	ADU75979 Thrombopo
57	56	100.0	14	9	ADV44319 Agonist T
58	56	100.0	14	9	AEI12792 TPO mimet
59	56	100.0	15	2	AAW35416 Thrombopo
60	56	100.0	15	2	AAW36784 Thrombopo
61	56	100.0	15	2	AAW36780 Thrombopo
62	56	100.0	15	2	AAW36776 Thrombopo
63	56	100.0	15	2	AAW66717 Peptide c
64	56	100.0	15	2	AAW66718 Peptide c
65	56	100.0	15	2	AAW66714 Peptide c
66	56	100.0	15	2	AAW66721 Peptide c
67	56	100.0	15	2	AAW66712 Peptide c
68	56	100.0	15	3	AAW20684 Thrombocy
69	56	100.0	15	4	AAU26022 Human thr
70	56	100.0	15	4	AAU25996 Human thr
71	56	100.0	15	4	AAU26026 Human thr
72	56	100.0	15	4	AAU26011 Human thr
73	56	100.0	15	4	AAU26020 Human thr
74	56	100.0	15	4	AAU25831 Human thr
75	56	100.0	15	4	AAU26023 Human thr
76	56	100.0	15	4	AAU26007 Human thr
77	56	100.0	15	5	ABP51670 Thrombopo
78	56	100.0	15	7	ABR62908 Thrombopo
79	56	100.0	15	8	ADM72485 TPO mimet
80	56	100.0	15	8	ADM72479 TPO mimet
81	56	100.0	15	8	ADM72502 TPO mimet
82	56	100.0	15	8	ADM72492 TPO mimet
83	56	100.0	15	8	ADM72478 TPO mimet
84	56	100.0	15	8	ADM72533 TPO mimet
85	56	100.0	15	8	ADM72496 TPO mimet
86	56	100.0	15	8	ADM72490 TPO mimet
87	56	100.0	15	8	ADM72486 TPO mimet
88	56	100.0	15	8	ADM72491 TPO mimet
89	56	100.0	15	8	ADM72522 TPO mimet
90	56	100.0	15	8	ADM72523 TPO mimet
91	56	100.0	15	8	ADM72493 TPO mimet
92	56	100.0	15	8	ADM72482 TPO mimet
93	56	100.0	15	8	ADM72494 TPO mimet
94	56	100.0	15	8	ADQ16585 TPO mimet
95	56	100.0	15	8	ADT92483 Modified
96	56	100.0	15	9	ADU70209 Thrombopo
97	56	100.0	15	9	ADU75981 Peptide-b

98	56	100.0	15	9	ADV44320	Adv44320 Agonist T	171	56	100.0	18	9	ASB12939	Aeb12939 TPO mimet
99	56	100.0	15	9	AEB12793	Aeb12793 TPO mimet	172	56	100.0	18	9	AEB12832	Aeb12832 TPO mimet
100	56	100.0	16	2	AAW19534	Aaw19534 Thrombopo	173	56	100.0	18	9	AEB12852	Aeb12852 TPO mimet
101	56	100.0	16	2	AAW33035	Aaw33035 Thrombopo	174	56	100.0	18	9	AEB12816	Aeb12816 TPO mimet
102	56	100.0	16	2	AAW36775	Aaw36775 Thrombopo	175	56	100.0	18	9	AEB12818	Aeb12818 TPO mimet
103	56	100.0	16	2	AAW36771	Aaw36771 Thrombopo	176	56	100.0	18	9	AEB12828	Aeb12828 TPO mimet
104	56	100.0	16	2	AAW66709	Aaw66709 Peptide c	177	56	100.0	18	9	AEB12836	Aeb12836 TPO mimet
105	56	100.0	16	2	AAW66713	Aaw66713 Peptide c	178	56	100.0	18	9	AEB12822	Aeb12822 TPO mimet
106	56	100.0	16	2	AAW66733	Aaw66733 Peptide c	179	56	100.0	18	9	AEB12824	Aeb12824 TPO mimet
107	56	100.0	16	2	AAW66716	Aaw66716 Peptide c	180	56	100.0	18	9	AEB12838	Aeb12838 TPO mimet
108	56	100.0	16	4	AAU26021	Aau26021 Human thr	181	56	100.0	19	5	ABB73391	Abb73391 TPO-mimet
109	56	100.0	16	4	AAU26005	Aau26005 Human thr	182	56	100.0	19	5	ABB73390	Abb73390 TPO-mimet
110	56	100.0	16	4	AAU26043	Aau26043 Human thr	183	56	100.0	20	3	AAB18003	Aab18003 Fc-TMP pe
111	56	100.0	16	4	AAU25832	Aau25832 Human thr	184	56	100.0	20	3	AAB17929	Aab17929 TPO-mimet
112	56	100.0	16	8	ADM72532	Adm72532 TPO mimet	185	56	100.0	20	5	ABB73403	Abb73403 TPO mimet
113	56	100.0	16	8	ADM72484	Adm72484 TPO mimet	186	56	100.0	21	7	ADN59687	Adn59687 Thrombopo
114	56	100.0	18	3	AAB16957	Aab16957 PEGylated	187	56	100.0	22	7	ADN59830	Adn59830 TMP pepti
115	56	100.0	18	3	AAB16956	Aab16956 PEGylated	188	56	100.0	22	7	ADN59819	Adn59819 TMP pepti
116	56	100.0	18	5	ABP51687	Abp51687 TPO mimet	189	56	100.0	22	8	ADQ16714	Adq16714 Immunoglo
117	56	100.0	18	5	ABP51689	Abp51689 TPO mimet	190	56	100.0	22	8	ADQ16713	Adq16713 Immunoglo
118	56	100.0	18	5	ABP51688	Abp51688 TPO mimet	191	56	100.0	22	8	ADQ16709	Adq16709 Immunoglo
119	56	100.0	18	5	ABP51677	Abp51677 TPO mimet	192	56	100.0	22	8	ADQ16706	Adq16706 Immunoglo
120	56	100.0	18	5	ABP51686	Abp51686 TPO mimet	193	56	100.0	22	8	ADQ16699	Adq16699 TPO mimet
121	56	100.0	18	5	ABP51674	Abp51674 TPO mimet	194	56	100.0	22	8	ADQ16712	Adq16712 Immunoglo
122	56	100.0	18	5	ABP51693	Abp51693 TPO mimet	195	56	100.0	22	8	ADQ16707	Adq16707 Immunoglo
123	56	100.0	18	5	ABP51684	Abp51684 TPO mimet	196	56	100.0	22	8	ADQ16711	Adq16711 Immunoglo
124	56	100.0	18	5	ABP51683	Abp51683 TPO mimet	197	56	100.0	22	8	ADQ16708	Adq16708 Immunoglo
125	56	100.0	18	5	ABP51685	Abp51685 TPO mimet	198	56	100.0	22	8	ADQ16710	Adq16710 Immunoglo
126	56	100.0	18	5	ABP51691	Abp51691 TPO mimet	199	56	100.0	22	9	ADV44435	Adv44435 Modified
127	56	100.0	18	5	ABP51673	Abp51673 TPO mimet	200	56	100.0	22	9	ADV44449	Adv44449 Anti-teta
128	56	100.0	18	5	ABP51690	Abp51690 TPO mimet	201	56	100.0	22	9	ADV44443	Adv44443 Anti-teta
129	56	100.0	18	5	ABP51675	Abp51675 TPO mimet	202	56	100.0	22	9	ADV44444	Adv44444 Anti-teta
130	56	100.0	18	5	ABP51692	Abp51692 TPO mimet	203	56	100.0	22	9	ADV44448	Adv44448 Anti-teta
131	56	100.0	18	7	ADN59663	Adn59663 Thrombopo	204	56	100.0	22	9	ADV44442	Adv44442 Anti-teta
132	56	100.0	18	7	ADN59812	Adn59812 Thrombopo	205	56	100.0	22	9	ADV44447	Adv44447 Anti-teta
133	56	100.0	18	8	ADQ16611	Adq16611 TPO mimet	206	56	100.0	22	9	ADV44445	Adv44445 Anti-teta
134	56	100.0	18	8	ADQ16619	Adq16619 TPO mimet	207	56	100.0	22	9	ADV44446	Adv44446 Anti-teta
135	56	100.0	18	8	ADQ16621	Adq16621 TPO mimet	208	56	100.0	22	9	AEB12918	Aeb12918 TPO mimet
136	56	100.0	18	8	ADQ16641	Adq16641 TPO mimet	209	56	100.0	22	9	AEB12917	Aeb12917 TPO mimet
137	56	100.0	18	8	ADQ16646	Adq16646 TPO mimet	210	56	100.0	22	9	AEB12919	Aeb12919 TPO mimet
138	56	100.0	18	8	ADQ16607	Adq16607 TPO mimet	211	56	100.0	22	9	AEB12986	Aeb12986 Tt antibo
139	56	100.0	18	8	ADQ16615	Adq16615 TPO mimet	212	56	100.0	22	9	AEB12986	Aeb12986 Tt antibo
140	56	100.0	18	8	ADQ16627	Adq16627 TPO mimet	213	56	100.0	22	9	AEB12920	Aeb12920 TPO mimet
141	56	100.0	18	8	ADQ16625	Adq16625 TPO mimet	214	56	100.0	22	9	AEB12916	Aeb12916 TPO mimet
142	56	100.0	18	8	ADQ16617	Adq16617 TPO mimet	215	56	100.0	22	9	AEB12914	Aeb12914 TPO mimet
143	56	100.0	18	8	ADQ16629	Adq16629 TPO mimet	216	56	100.0	22	9	AEB12921	Aeb12921 TPO mimet
144	56	100.0	18	8	ADQ16613	Adq16613 TPO mimet	217	56	100.0	22	9	AEB12907	Aeb12907 Tt antibo
145	56	100.0	18	8	ADQ16623	Adq16623 TPO mimet	218	56	100.0	22	9	AEB12915	Aeb12915 TPO mimet
146	56	100.0	18	8	ADQ16605	Adq16605 TPO mimet	219	56	100.0	25	7	ADN59708	Adn59708 Thrombopo
147	56	100.0	18	8	ADQ16609	Adq16609 TPO mimet	220	56	100.0	28	3	AAB17285	Aab17285 TPO-mimet
148	56	100.0	18	9	ADV44355	Adv44355 Agonist T	221	56	100.0	28	5	ABP51682	Abp51682 TPO mimet
149	56	100.0	18	9	ADV44345	Adv44345 Agonist T	222	56	100.0	28	7	ADJ73013	Adj73013 TPO mimet
150	56	100.0	18	9	ADV44351	Adv44351 Agonist T	223	56	100.0	28	8	ADJ52648	Adj52648 CH1 delet
151	56	100.0	18	9	ADV44357	Adv44357 Agonist T	224	56	100.0	28	8	ADJ51609	Adj51609 CH1 delet
152	56	100.0	18	9	ADV44349	Adv44349 Agonist T	225	56	100.0	28	8	ADQ16636	Adq16636 Tetanus t
153	56	100.0	18	9	ADV44343	Adv44343 Agonist T	226	56	100.0	28	9	ADV44374	Adv44374 Modified
154	56	100.0	18	9	ADV44347	Adv44347 Agonist T	227	56	100.0	28	9	AEB12847	Aeb12847 Antibody
155	56	100.0	18	9	ADV44384	Adv44384 Agonist T	228	56	100.0	29	3	AAB16971	Aab16971 TPO-mimet
156	56	100.0	18	9	ADV44353	Adv44353 Agonist T	229	56	100.0	29	3	AAB16975	Aab16975 TPO-mimet
157	56	100.0	18	9	ADV44363	Adv44363 Agonist T	230	56	100.0	29	3	AAB16976	Aab16976 TPO-mimet
158	56	100.0	18	9	ADV44361	Adv44361 Agonist T	231	56	100.0	29	3	AAB17286	Aab17286 TPO-mimet
159	56	100.0	18	9	ADV44379	Adv44379 Agonist T	232	56	100.0	29	3	AAB16970	Aab16970 TPO-mimet
160	56	100.0	18	9	ADV44344	Adv44344 Agonist T	233	56	100.0	29	5	ABB72862	Abb72862 TPO mimet
161	56	100.0	18	9	ADV44359	Adv44359 Agonist T	234	56	100.0	29	5	ABB72857	Abb72857 TPO mimet
162	56	100.0	18	9	ADV44365	Adv44365 Agonist T	235	56	100.0	29	5	ABB72861	Abb72861 TPO mimet
163	56	100.0	18	9	ADV44367	Adv44367 Agonist T	236	56	100.0	29	5	ABB72856	Abb72856 TPO mimet
164	56	100.0	18	9	ADV44468	Adv44468 Agonist T	237	56	100.0	29	7	ADJ73011	Adj73011 TPO mimet
165	56	100.0	18	9	AEB12826	Aeb12826 TPO mimet	238	56	100.0	29	7	ADJ73007	Adj73007 TPO mimet
166	56	100.0	18	9	AEB12820	Aeb12820 TPO mimet	239	56	100.0	29	7	ADJ73006	Adj73006 TPO mimet
167	56	100.0	18	9	AEB12840	Aeb12840 TPO mimet	240	56	100.0	29	8	ADJ52642	Adj52642 CH1 delet
168	56	100.0	18	9	AEB12857	Aeb12857 Antibody	241	56	100.0	29	8	ADJ52646	Adj52646 CH1 delet
169	56	100.0	18	9	AEB12834	Aeb12834 TPO mimet	242	56	100.0	29	8	ADJ52641	Adj52641 CH1 delet
170	56	100.0	18	9	AEB12830	Aeb12830 TPO mimet	243	56	100.0	29	8	ADJ51603	Adj51603 CH1 delet

244 56 100.0 29 8 ADJ51602 Adj51602 CH1 delet 317
245 56 100.0 29 8 ADJ51607 Adj51607 CH1 delet 318
246 56 100.0 30 3 AAB17287 Aab17287 TPO-mimet 319
247 56 100.0 31 3 AAB17288 Aab17288 TPO-mimet 320
248 56 100.0 31 3 AAB16974 Aab16974 TPO-mimet 321
249 56 100.0 31 3 AAB16973 Aab16973 TPO-mimet 322
250 56 100.0 31 5 ABB72860 Aab16973 TPO-mimet 323
251 56 100.0 31 5 ABB72859 Aab16973 TPO-mimet 324
252 56 100.0 31 7 ADJ73009 Adj73009 TPO-mimet 325
253 56 100.0 31 7 ADJ73010 Adj73010 TPO-mimet 326
254 56 100.0 31 8 ADJ52644 Adj52644 CH1 delet 327
255 56 100.0 31 8 ADJ52645 Adj52645 CH1 delet 328
256 56 100.0 31 8 ADJ51606 Adj51606 CH1 delet 329
257 56 100.0 31 8 ADJ51605 Adj51605 CH1 delet 330
258 56 100.0 32 3 AAY96520 Aay96520 Thrombopo 331
259 56 100.0 32 3 AAB17289 Aab17289 TPO-mimet 332
260 56 100.0 32 3 AAB17297 Aab17297 TPO-mimet 333
261 56 100.0 33 3 AAB17290 Aab17290 TPO-mimet 334
262 56 100.0 34 3 AAY96527 Aay96527 Thrombopo 335
263 56 100.0 34 3 AAB17291 Aab17291 TPO-mimet 336
264 56 100.0 35 3 AAB17292 Aab17292 TPO-mimet 337
265 56 100.0 36 3 AAY96525 Aay96525 Thrombopo 338
266 56 100.0 36 3 AAY96523 Aay96523 Thrombopo 339
267 56 100.0 36 3 AAY96524 Aay96524 Thrombopo 340
268 56 100.0 36 3 AAY96526 Aay96526 Thrombopo 341
269 56 100.0 36 3 AAB17307 Aab17307 TPO-mimet 342
270 56 100.0 36 3 AAB17293 Aab17293 TPO-mimet 343
271 56 100.0 36 3 AAB17303 Aab17303 TPO-mimet 344
272 56 100.0 36 3 AAB16963 Aab16963 TPO-mimet 345
273 56 100.0 36 3 AAB17301 Aab17301 TPO-mimet 346
274 56 100.0 36 3 AAB17306 Aab17306 TPO-mimet 347
275 56 100.0 36 5 ABB72403 Aab17306 TPO-mimet 348
276 56 100.0 37 3 AAB17294 Aab17294 TPO-mimet 349
277 56 100.0 38 3 AAB17295 Aab17295 TPO-mimet 350
278 56 100.0 39 3 AAB17304 Aab17304 TPO-mimet 351
279 56 100.0 39 3 AAB17305 Aab17305 TPO-mimet 352
280 56 100.0 40 3 AAB17302 Aab17302 TPO-mimet 353
281 56 100.0 41 3 AAY96528 Aay96528 Thrombopo 354
282 56 100.0 41 5 ABB73389 Aab73389 TPO-mimet 355
283 56 100.0 41 5 ABB73388 Aab73388 TPO-mimet 356
284 56 100.0 42 3 AAY96530 Aay96530 Thrombopo 357
285 56 100.0 42 3 AAB17296 Aab17296 TPO-mimet 358
286 56 100.0 42 3 AAB17308 Aab17308 Synthetic 359
287 56 100.0 42 3 AAB17282 Aab17282 TPO-mimet 360
288 56 100.0 42 3 AAB17281 Aab17281 TPO-mimet 361
289 56 100.0 42 5 ABB73404 Aab73404 TMP-TMP g 362
290 56 100.0 43 7 ADN59759 Adn59759 Peptide-v 363
291 56 100.0 60 3 AAB17311 Aab17311 Synthetic 364
292 56 100.0 60 5 ABB73405 Aab73405 TMP-TMP g 365
293 56 100.0 122 9 ADV44474 Adv44474 Anti-teta 366
294 56 100.0 122 9 AEB12946 Aeb12946 Antibody 367
295 56 100.0 128 8 ADQ16705 Adq16705 Modified 368
296 56 100.0 128 9 ADV44466 Adv44466 Anti-teta 369
297 56 100.0 128 9 ADV44463 Adv44463 Anti-teta 370
298 56 100.0 128 9 ADV44467 Adv44467 Anti-teta 371
299 56 100.0 128 9 ADV44465 Adv44465 Anti-teta 372
300 56 100.0 128 9 ADV44441 Adv44441 pAX116 va 373
301 56 100.0 128 9 ADV44464 Adv44464 Anti-teta 374
302 56 100.0 128 9 AEB12934 Aeb12934 Antibody 375
303 56 100.0 128 9 AEB12913 Aeb12913 Antibody 376
304 56 100.0 128 9 AEB12913 Aeb12913 Antibody 377
305 56 100.0 128 9 AEB12936 Aeb12936 Antibody 378
306 56 100.0 128 9 AEB12937 Aeb12937 Antibody 379
307 56 100.0 128 9 AEB12938 Aeb12938 Antibody 380
308 56 100.0 129 6 ABG71751 Abg71751 Antibody 381
309 56 100.0 131 6 ABG71753 Abg71753 Antibody 382
310 56 100.0 132 9 ADV44473 Adv44473 Anti-teta 383
311 56 100.0 133 6 ABG71752 Abg71752 Antibody 384
312 56 100.0 135 6 ABG71749 Abg71749 Antibody 385
313 56 100.0 143 6 ABG71750 Abg71750 Antibody 386
314 56 100.0 144 6 ABG71748 Abg71748 Antibody 387
315 56 100.0 150 9 AEB12945 Aeb12945 Antibody 388
316 56 100.0 225 8 ADQ16704 Adq16704 Modified 389

56 100.0 234 9 AEB12912 Aeb12912 Antibody 317
56 100.0 247 3 AAB16958 Aab16958 Fc-TMP pr 318
56 100.0 247 3 AAB16961 Aab16961 TMP-Fc pr 319
56 100.0 247 5 ABB73411 Abb73411 Fc-TPO mi 320
56 100.0 247 5 ABB73414 Abb73414 TMP-Fc am 321
56 100.0 249 9 ADV44440 Adv44440 pAX116 va 322
56 100.0 268 3 AAB16959 Aab16959 Fc-TMP-TM 323
56 100.0 268 5 ABB73412 Abb73412 Fc-TMP-TM 324
56 100.0 269 3 AAY96531 Aay96531 Human IGG 325
56 100.0 269 3 AAB16960 Aab16960 TMP-TMP-F 326
56 100.0 269 5 ABB73413 Abb73413 TMP-TMP-F 327
56 100.0 282 9 AEB12930 Aebi2930 Antibody 328
56 100.0 459 9 ADV44459 Adv44459 Anti-teta 329
56 100.0 472 5 ABP51695 Abp51695 5G1.1-TPO 330
56 100.0 472 8 ADQ16647 Adq16647 Immunoglo 331
56 100.0 472 9 ADV44385 Adv44385 5G1.1 hea 332
56 100.0 472 9 AEB12858 Aebi2858 Antibody 333
53 94.6 14 3 AAB16969 Aab16969 TPO-mimet 334
53 94.6 14 5 ABB72855 Abb72855 TPO-mimet 335
53 94.6 14 7 ADJ73005 Adj73005 TPO-mimet 336
53 94.6 14 8 ADJ52640 Adj52640 CH1 delet 337
53 94.6 14 8 ADJ51601 Adj51601 CH1 delet 338
53 94.6 15 2 AAW66719 Aaw66719 Peptide c 339
52 92.9 13 4 AAU26018 Aau26018 Human thr 340
52 92.9 13 4 AAU26035 Aau26035 Human thr 341
52 92.9 13 8 ADM72525 Adm72525 TPO-mimet 342
52 92.9 13 8 ADM72488 Adm72488 TPO-mimet 343
52 92.9 13 8 ADM72489 Adm72489 TPO-mimet 344
52 92.9 13 8 ADM72524 Adm72524 TPO-mimet 345
52 92.9 14 2 AAW36773 Aaw36773 Thrombopo 346
52 92.9 14 2 AAW66715 Aaw66715 Peptide c 347
52 92.9 14 2 AAW66730 Aaw66730 Peptide c 348
52 92.9 14 4 AAU26019 Aau26019 Human thr 349
52 92.9 14 4 AAU26036 Aau26036 Human thr 350
52 92.9 14 4 AAU26037 Aau26037 Human thr 351
52 92.9 14 8 ADM72526 Adm72526 TPO-mimet 352
52 92.9 14 8 ADM72527 Adm72527 TPO-mimet 353
52 92.9 15 2 AAW66731 Aaw66731 Peptide c 354
52 92.9 15 4 AAU26038 Aau26038 Human thr 355
52 92.9 18 7 ADN59660 Adn59660 Thrombopo 356
52 92.9 18 7 ADN59815 Adn59815 Thrombopo 357
52 92.9 18 7 ADN59653 Adn59653 Thrombopo 358
52 92.9 18 7 ADN59670 Adn59670 Thrombopo 359
52 92.9 19 2 AAW09494 Aaw09494 Thrombopo 360
52 92.9 19 2 AAW33028 Aaw33028 Thrombopo 361
52 92.9 19 2 AAW36645 Aaw36645 Thrombopo 362
52 92.9 19 3 AAB17022 Aab17022 TPO-mimet 363
52 92.9 19 4 AAU25864 Aau25864 Human thr 364
52 92.9 19 4 AAU25825 Aau25825 Human thr 365
52 92.9 19 5 ABB72908 Abb72908 TPO-mimet 366
52 92.9 19 7 ADJ73060 Adj73060 TPO-mimet 367
52 92.9 19 8 ADJ52695 Adj52695 CH1 delet 368
52 92.9 19 8 ADJ51656 Adj51656 CH1 delet 369
52 92.9 22 7 ADN59820 Adn59820 TMP pepti 370
52 92.9 22 7 ADN59827 Adn59827 TMP pepti 371
52 92.9 22 7 ADN59837 Adn59837 TMP pepti 372
52 92.9 25 7 ADN59689 Adn59689 Thrombopo 373
52 92.9 25 7 ADN59722 Adn59722 Thrombopo 374
52 92.9 25 7 ADN59702 Adn59702 Thrombopo 375
52 92.9 25 7 ADN59751 Adn59751 Peptide-v 376
52 92.9 42 7 ADN59678 Adn59678 Thrombopo 377
51 91.1 18 7 ADN59678 Adn59678 Thrombopo 378
51 91.1 19 2 AAW09491 Aaw09491 Thrombopo 379
51 91.1 19 2 AAW09493 Aaw09493 Thrombopo 380
51 91.1 19 2 AAW36644 Aaw36644 Thrombopo 381
51 91.1 19 2 AAW35418 Aaw35418 Thrombopo 382
51 91.1 19 2 AAW36642 Aaw36642 Thrombopo 383
51 91.1 19 4 AAU25861 Aau25861 Human thr 384
51 91.1 19 4 AAU25863 Aau25863 Human thr 385
51 91.1 19 4 AAU25998 Aau25998 Human thr 386
51 91.1 25 7 ADN59736 Adn59736 Thrombopo 387
50 89.3 10 2 AAW09469 Aaw09469 Thrombopo 388
50 89.3 10 2 AAW36621 Aaw36621 Thrombopo 389

390	50	89.3	10	2	AAW36772	Aaw36772	Thrombopo	463	49	87.5	19	4	AAU25870	Aau25870	Human thr
391	50	89.3	10	3	AAB17005	Aab17005	TPO-mimet	464	49	87.5	19	4	AAU25821	Aau25821	Human thr
392	50	89.3	10	4	AAU25839	Aau25839	Human thr	465	49	87.5	19	5	ABB72907	Abb72907	TPO mimet
393	50	89.3	10	4	AAU25840	Aau25840	Human thr	466	49	87.5	19	7	ADJ73059	Adj73059	TPO mimet
394	50	89.3	10	5	ABB72891	Abb72891	TPO mimet	467	49	87.5	19	8	ADJ52694	Adj52694	CH1 delet
395	50	89.3	10	7	ADJ73042	Adj73042	TPO mimet	468	49	87.5	19	8	ADJ51655	Adj51655	CH1 delet
396	50	89.3	10	8	ADJ52677	Adj52677	CH1 delet	469	49	87.5	22	7	ADN59826	Adn59826	TMP pepti
397	50	89.3	10	8	ADJ51638	Adj51638	CH1 delet	470	49	87.5	22	7	ADN59836	Adn59836	TMP pepti
398	50	89.3	14	8	ADM72499	Adm72499	TPO mimet	471	49	87.5	22	7	ADN59821	Adn59821	TMP pepti
399	50	89.3	15	4	AAU26024	Aau26024	Human thr	472	49	87.5	22	7	ADN59838	Adn59838	TMP pepti
400	50	89.3	15	8	ADM72498	Adm72498	TPO mimet	473	49	87.5	22	7	ADN59839	Adn59839	TMP pepti
401	50	89.3	18	7	ADN59664	Adn59664	Thrombopo	474	49	87.5	22	7	ADN59834	Adn59834	TMP pepti
402	50	89.3	18	7	ADN59668	Adn59668	Thrombopo	475	49	87.5	23	7	ADN59795	Adn59795	Peptide-v
403	50	89.3	18	7	ADN59666	Adn59666	Thrombopo	476	49	87.5	23	7	ADN59777	Adn59777	Peptide-v
404	50	89.3	22	7	ADN59835	Adn59835	TMP pepti	477	49	87.5	25	7	ADN59742	Adn59742	Thrombopo
405	50	89.3	22	7	ADN59831	Adn59831	TMP pepti	478	49	87.5	25	7	ADN59716	Adn59716	Thrombopo
406	50	89.3	22	7	ADN59833	Adn59833	TMP pepti	479	49	87.5	25	7	ADN59724	Adn59724	Thrombopo
407	50	89.3	25	7	ADN59710	Adn59710	Thrombopo	480	49	87.5	25	7	ADN59691	Adn59691	Thrombopo
408	50	89.3	25	7	ADN59718	Adn59718	Thrombopo	481	49	87.5	25	7	ADN59700	Adn59700	Thrombopo
409	50	89.3	25	7	ADN59714	Adn59714	Thrombopo	482	49	87.5	25	7	ADN59720	Adn59720	Thrombopo
410	49	87.5	10	2	AAW09472	Aaw09472	Thrombopo	483	49	87.5	25	7	ADN59744	Adn59744	Thrombopo
411	49	87.5	10	2	AAW36623	Aaw36623	Thrombopo	484	49	87.5	36	7	ADN59765	Adn59765	Peptide-v
412	49	87.5	10	4	AAU25842	Aau25842	Human thr	485	49	87.5	41	7	ADN59771	Adn59771	Peptide-v
413	49	87.5	13	3	AAB17015	Aab17015	TPO-mimet	486	49	87.5	42	7	ADN59818	Adn59818	Peptide-
414	49	87.5	13	5	ABB72901	Abb72901	TPO mimet	487	49	87.5	43	7	ADN59752	Adn59752	Peptide-v
415	49	87.5	13	7	ADJ73054	Adj73054	TPO mimet	488	49	87.5	46	7	ADN59789	Adn59789	Peptide-v
416	49	87.5	13	7	ADJ73052	Adj73052	TPO mimet	489	49	87.5	46	7	ADN59757	Adn59757	Peptide-v
417	49	87.5	13	7	ADJ73056	Adj73056	TPO mimet	490	49	87.5	46	7	ADN59783	Adn59783	Peptide-v
418	49	87.5	13	7	ADJ73053	Adj73053	TPO mimet	491	48	85.7	12	3	AAB17309	Aab17309	Synthetic
419	49	87.5	13	7	ADJ73055	Adj73055	TPO mimet	492	48	85.7	18	2	AAW09499	Aaw09499	Thrombopo
420	49	87.5	13	8	ADJ52687	Adj52687	CH1 delet	493	48	85.7	18	2	AAW09459	Aaw09459	Thrombopo
421	49	87.5	13	8	ADJ51648	Adj51648	CH1 delet	494	48	85.7	18	2	AAW36650	Aaw36650	Thrombopo
422	49	87.5	14	3	AAB17017	Aab17017	TPO-mimet	495	48	85.7	18	2	AAW33026	Aaw33026	Thrombopo
423	49	87.5	14	3	AAB17016	Aab17016	TPO-mimet	496	48	85.7	18	3	AAB17024	Aab17024	TPO-mimet
424	49	87.5	14	5	ABB72903	Abb72903	TPO mimet	497	48	85.7	18	4	AAU25869	Aau25869	Human thr
425	49	87.5	14	5	ABB72902	Abb72902	TPO mimet	498	48	85.7	18	4	AAU25823	Aau25823	Human thr
426	49	87.5	14	8	ADJ52689	Adj52689	CH1 delet	499	48	85.7	18	5	ABB72910	Abb72910	TPO mimet
427	49	87.5	14	8	ADJ52688	Adj52688	CH1 delet	500	48	85.7	18	7	ADJ73062	Adj73062	TPO mimet
428	49	87.5	14	8	ADJ51650	Adj51650	CH1 delet	501	48	85.7	18	7	ADN59675	Adn59675	Thrombopo
429	49	87.5	14	8	ADJ51649	Adj51649	CH1 delet	502	48	85.7	18	7	ADN59680	Adn59680	Thrombopo
430	49	87.5	15	3	AAB17018	Aab17018	TPO-mimet	503	48	85.7	18	7	ADN59661	Adn59661	Thrombopo
431	49	87.5	15	5	ABB72904	Abb72904	TPO mimet	504	48	85.7	18	7	ADN59657	Adn59657	Thrombopo
432	49	87.5	15	8	ADJ52691	Adj52691	CH1 delet	505	48	85.7	18	8	ADJ52697	Adj52697	CH1 delet
433	49	87.5	15	8	ADJ52690	Adj52690	CH1 delet	506	48	85.7	18	8	ADJ51658	Adj51658	CH1 delet
434	49	87.5	15	8	ADJ51652	Adj51652	CH1 delet	507	48	85.7	22	7	ADN59824	Adn59824	TMP pepti
435	49	87.5	15	8	ADJ51651	Adj51651	CH1 delet	508	48	85.7	22	7	ADN59828	Adn59828	TMP pepti
436	49	87.5	18	2	AAW09460	Aaw09460	Thrombopo	509	48	85.7	23	7	ADN59775	Adn59775	Peptide-v
437	49	87.5	18	2	AAW09498	Aaw09498	Thrombopo	510	48	85.7	23	7	ADN59793	Adn59793	Peptide-v
438	49	87.5	18	2	AAW36649	Aaw36649	Thrombopo	511	48	85.7	25	7	ADN59696	Adn59696	Thrombopo
439	49	87.5	18	2	AAW33027	Aaw33027	Thrombopo	512	48	85.7	25	7	ADN59704	Adn59704	Thrombopo
440	49	87.5	18	2	AAW36652	Aaw36652	Thrombopo	513	48	85.7	25	7	ADN59730	Adn59730	Thrombopo
441	49	87.5	18	3	AAB17026	Aab17026	TPO-mimet	514	48	85.7	25	7	ADN59740	Adn59740	Thrombopo
442	49	87.5	18	4	AAU25868	Aau25868	Human thr	515	48	85.7	36	7	ADN59763	Adn59763	Peptide-v
443	49	87.5	18	4	AAU25824	Aau25824	Human thr	516	48	85.7	40	7	ADN59753	Adn59753	Peptide-v
444	49	87.5	18	4	AAU25871	Aau25871	Human thr	517	48	85.7	41	7	ADN59769	Adn59769	Peptide-v
445	49	87.5	18	5	ABB72912	Abb72912	TPO mimet	518	48	85.7	44	7	ADN59817	Adn59817	Peptide-
446	49	87.5	18	7	ADJ73064	Adj73064	TPO mimet	519	48	85.7	46	7	ADN59781	Adn59781	Peptide-v
447	49	87.5	18	7	ADN59681	Adn59681	Thrombopo	520	48	85.7	46	7	ADN59787	Adn59787	Peptide-v
448	49	87.5	18	7	ADN59672	Adn59672	Thrombopo	521	47	83.9	8	2	AAW33037	Aaw33037	Thrombopo
449	49	87.5	18	7	ADN59654	Adn59654	Thrombopo	522	47	83.9	8	4	AAU25982	Aau25982	Human thr
450	49	87.5	18	7	ADN59671	Adn59671	Thrombopo	523	47	83.9	8	5	ABP51678	Abp51678	TPO mimet
451	49	87.5	18	7	ADN59659	Adn59659	Thrombopo	524	47	83.9	8	8	ADQ16692	Adq16692	TPO mimet
452	49	87.5	18	7	ADN59669	Adn59669	Thrombopo	525	47	83.9	8	9	ADV44430	Adv44430	Agonist T
453	49	87.5	18	7	ADN59667	Adn59667	Thrombopo	526	47	83.9	8	9	AEB12902	Aeb12902	TPO mimet
454	49	87.5	18	8	ADJ52699	Adj52699	CH1 delet	527	47	83.9	18	5	ABP51679	Abp51679	TPO mimet
455	49	87.5	18	8	ADJ51660	Adj51660	CH1 delet	528	47	83.9	18	8	ADQ16693	Adq16693	TPO mimet
456	49	87.5	19	2	AAW09457	Aaw09457	Thrombopo	529	47	83.9	18	9	ADV44431	Adv44431	Anti-teta
457	49	87.5	19	2	AAW09492	Aaw09492	Thrombopo	530	47	83.9	18	9	AEB12903	Aeb12903	Heavy cha
458	49	87.5	19	2	AAW36651	Aaw36651	Thrombopo	531	46	82.1	14	8	ADM72509	Adm72509	TPO mimet
459	49	87.5	19	2	AAW33024	Aaw33024	Thrombopo	532	46	82.1	15	2	AAW66723	Aaw66723	Peptide c
460	49	87.5	19	2	AAW36643	Aaw36643	Thrombopo	533	46	82.1	15	4	AAU26028	Aau26028	Human thr
461	49	87.5	19	3	AAB17021	Aab17021	TPO-mimet	534	46	82.1	15	8	ADM72508	Adm72508	TPO mimet
462	49	87.5	19	4	AAU25862	Aau25862	Human thr	535	46	82.1	18	7	ADN59658	Adn59658	Thrombopo

536	46	82.1	18	7	ADN59665	Adn59665	Thrombopo	609	45	80.4	14	4	AAU25866	Aau25866	Human thr
537	46	82.1	18	7	ADN59655	Adn59655	Thrombopo	610	45	80.4	14	4	AAU25988	Aau25988	Human thr
538	46	82.1	18	7	ADN59656	Adn59656	Thrombopo	611	45	80.4	14	5	ABB72900	Abb72900	TPO mimet
539	46	82.1	22	7	ADN59823	Adn59823	TMP pepti	612	45	80.4	14	7	ADJ73051	Adj73051	TPO mimet
540	46	82.1	22	7	ADN59822	Adn59822	TMP pepti	613	45	80.4	14	8	ADJ52686	Adj52686	CH1 delet
541	46	82.1	22	7	ADN59825	Adn59825	TMP pepti	614	45	80.4	14	8	ADJ51647	Adj51647	CH1 delet
542	46	82.1	22	7	ADN59832	Adn59832	TMP pepti	615	45	80.4	18	2	AAW09456	Aaw09456	Thrombopo
543	46	82.1	23	7	ADN59776	Adn59776	Peptide-v	616	45	80.4	18	2	AAW09487	Aaw09487	Thrombopo
544	46	82.1	23	7	ADN59792	Adn59792	Peptide-v	617	45	80.4	18	2	AAW33023	Aaw33023	Thrombopo
545	46	82.1	23	7	ADN59774	Adn59774	Peptide-v	618	45	80.4	18	2	AAW36638	Aaw36638	Thrombopo
546	46	82.1	23	7	ADN59794	Adn59794	Peptide-v	619	45	80.4	18	3	AAB17020	Aab17020	TPO-mimet
547	46	82.1	25	7	ADN59694	Adn59694	Thrombopo	620	45	80.4	18	4	AAU25857	Aau25857	Human thr
548	46	82.1	25	7	ADN59692	Adn59692	Thrombopo	621	45	80.4	18	4	AAU25820	Aau25820	Human thr
549	46	82.1	25	7	ADN59698	Adn59698	Thrombopo	622	45	80.4	18	5	ABB72906	Abb72906	TPO mimet
550	46	82.1	25	7	ADN59712	Adn59712	Thrombopo	623	45	80.4	18	7	ADJ73058	Adj73058	TPO mimet
551	46	82.1	25	7	ADN59755	Adn59755	Peptide-v	624	45	80.4	18	8	ADJ52693	Adj52693	CH1 delet
552	46	82.1	36	7	ADN59762	Adn59762	Peptide-v	625	45	80.4	18	8	ADJ51654	Adj51654	CH1 delet
553	46	82.1	36	7	ADN59764	Adn59764	Peptide-v	626	45	80.4	19	2	AAW09458	Aaw09458	Thrombopo
554	46	82.1	41	7	ADN59768	Adn59768	Peptide-v	627	45	80.4	19	2	AAW33025	Aaw33025	Thrombopo
555	46	82.1	41	7	ADN59770	Adn59770	Peptide-v	628	45	80.4	19	4	AAU25822	Aau25822	Human thr
556	46	82.1	43	7	ADN59761	Adn59761	Peptide-v	629	45	80.4	25	4	AAU26042	Aau26042	Human thr
557	46	82.1	46	7	ADN59780	Adn59780	Peptide-v	630	45	80.4	25	8	ADM72531	Adm72531	TPO mimet
558	46	82.1	46	7	ADN59788	Adn59788	Peptide-v	631	44	78.6	12	8	ADM72530	Adm72530	TPO mimet
559	46	82.1	46	7	ADN59786	Adn59786	Peptide-v	632	44	78.6	13	4	AAU26041	Aau26041	Human thr
560	46	82.1	46	7	ADN59782	Adn59782	Peptide-v	633	44	78.6	13	4	AAU26039	Aau26039	Human thr
561	45	80.4	11	2	AAW36791	Aaw36791	Thrombopo	634	44	78.6	13	8	ADM72529	Adm72529	TPO mimet
562	45	80.4	12	2	AAW35423	Aaw35423	Thrombopo	635	44	78.6	13	8	ADM72528	Adm72528	TPO mimet
563	45	80.4	12	4	AAU26000	Aau26000	Human thr	636	44	78.6	14	2	AAW66732	Aaw66732	Peptide c
564	45	80.4	13	2	AAW09467	Aaw09467	Thrombopo	637	44	78.6	14	4	AAU26040	Aau26040	Human thr
565	45	80.4	13	2	AAW35399	Aaw35399	Thrombopo	638	44	78.6	18	2	AAW09497	Aaw09497	Thrombopo
566	45	80.4	13	2	AAW35404	Aaw35404	Thrombopo	639	44	78.6	18	2	AAW36653	Aaw36653	Thrombopo
567	45	80.4	13	2	AAW35417	Aaw35417	Thrombopo	640	44	78.6	18	2	AAW36648	Aaw36648	Thrombopo
568	45	80.4	13	2	AAW33033	Aaw33033	Thrombopo	641	44	78.6	18	3	AAB17025	Aab17025	TPO-mimet
569	45	80.4	13	2	AAW35413	Aaw35413	Thrombopo	642	44	78.6	18	4	AAU25872	Aau25872	Human thr
570	45	80.4	13	2	AAW35406	Aaw35406	Thrombopo	643	44	78.6	18	4	AAU25867	Aau25867	Human thr
571	45	80.4	13	2	AAW35405	Aaw35405	Thrombopo	644	44	78.6	18	5	ABB72911	Abb72911	TPO mimet
572	45	80.4	13	2	AAW35422	Aaw35422	Thrombopo	645	44	78.6	18	7	ADJ73063	Adj73063	TPO mimet
573	45	80.4	13	2	AAW35397	Aaw35397	Thrombopo	646	44	78.6	18	7	ADN59674	Adn59674	Thrombopo
574	45	80.4	13	2	AAW36783	Aaw36783	Thrombopo	647	44	78.6	18	7	ADN59673	Adn59673	Thrombopo
575	45	80.4	13	4	AAU25994	Aau25994	Human thr	648	44	78.6	18	8	ADJ52698	Adj52698	CH1 delet
576	45	80.4	13	4	AAU25997	Aau25997	Human thr	649	44	78.6	18	8	ADJ51659	Adj51659	CH1 delet
577	45	80.4	13	4	AAU25991	Aau25991	Human thr	650	44	78.6	22	7	ADN59840	Adn59840	TMP pepti
578	45	80.4	13	4	AAU25990	Aau25990	Human thr	651	44	78.6	22	7	ADN59841	Adn59841	TMP pepti
579	45	80.4	13	4	AAU25984	Aau25984	Human thr	652	44	78.6	23	7	ADN59797	Adn59797	Peptide-v
580	45	80.4	14	2	AAW09466	Aaw09466	Thrombopo	653	44	78.6	23	7	ADN59778	Adn59778	Peptide-v
581	45	80.4	14	2	AAW09465	Aaw09465	Thrombopo	654	44	78.6	23	7	ADN59779	Adn59779	Peptide-v
582	45	80.4	14	2	AAW09462	Aaw09462	Thrombopo	655	44	78.6	23	7	ADN59796	Adn59796	Peptide-v
583	45	80.4	14	2	AAW09482	Aaw09482	Thrombopo	656	44	78.6	25	7	ADN59726	Adn59726	Thrombopo
584	45	80.4	14	2	AAW33031	Aaw33031	Thrombopo	657	44	78.6	25	7	ADN59728	Adn59728	Thrombopo
585	45	80.4	14	2	AAW35412	Aaw35412	Thrombopo	658	44	78.6	35	7	ADN59754	Adn59754	Peptide-v
586	45	80.4	14	2	AAW35407	Aaw35407	Thrombopo	659	44	78.6	36	7	ADN59766	Adn59766	Peptide-v
587	45	80.4	14	2	AAW35408	Aaw35408	Thrombopo	660	44	78.6	36	7	ADN59756	Adn59756	Peptide-v
588	45	80.4	14	2	AAW35398	Aaw35398	Thrombopo	661	44	78.6	36	7	ADN59767	Adn59767	Peptide-
589	45	80.4	14	2	AAW36633	Aaw36633	Thrombopo	662	44	78.6	41	7	ADN59816	Adn59816	Peptide-
590	45	80.4	14	2	AAW33029	Aaw33029	Thrombopo	663	44	78.6	41	7	ADN59772	Adn59772	Peptide-v
591	45	80.4	14	2	AAW35396	Aaw35396	Thrombopo	664	44	78.6	41	7	ADN59773	Adn59773	Peptide-v
592	45	80.4	14	2	AAW35401	Aaw35401	Thrombopo	665	44	78.6	46	7	ADN59790	Adn59790	Peptide-v
593	45	80.4	14	2	AAW35403	Aaw35403	Thrombopo	666	44	78.6	46	7	ADN59785	Adn59785	Peptide-v
594	45	80.4	14	2	AAW36647	Aaw36647	Thrombopo	667	44	78.6	46	7	ADN59784	Adn59784	Peptide-v
595	45	80.4	14	2	AAW35400	Aaw35400	Thrombopo	668	44	78.6	46	7	ADN59791	Adn59791	Peptide-v
596	45	80.4	14	2	AAW35402	Aaw35402	Thrombopo	669	44	78.6	75	7	ADN59758	Adn59758	Peptide-v
597	45	80.4	14	2	AAW33032	Aaw33032	Thrombopo	670	43	76.8	11	2	AAW35425	Aaw35425	Thrombopo
598	45	80.4	14	3	AAB17014	Aab17014	TPO-mimet	671	43	76.8	11	4	AAU26001	Aau26001	Human thr
599	45	80.4	14	4	AAU25987	Aau25987	Human thr	672	43	76.8	14	8	ADM72507	Adm72507	TPO mimet
600	45	80.4	14	4	AAU25826	Aau25826	Human thr	673	43	76.8	14	8	ADM72511	Adm72511	TPO mimet
601	45	80.4	14	4	AAU25993	Aau25993	Human thr	674	43	76.8	14	8	ADM72505	Adm72505	TPO mimet
602	45	80.4	14	4	AAU25852	Aau25852	Human thr	675	43	76.8	14	8	ADM72475	Adm72475	TPO mimet
603	45	80.4	14	4	AAU25989	Aau25989	Human thr	676	43	76.8	14	8	ADM72501	Adm72501	TPO mimet
604	45	80.4	14	4	AAU25983	Aau25983	Human thr	677	43	76.8	14	8	ADM72505	Adm72505	Thrombopo
605	45	80.4	14	4	AAU25985	Aau25985	Human thr	678	43	76.8	14	9	ADU70205	Adu70205	Thrombopo
606	45	80.4	14	4	AAU25995	Aau25995	Human thr	679	43	76.8	14	9	ADU75978	Adu75978	Thrombopo
607	45	80.4	14	4	AAU25992	Aau25992	Human thr	680	43	76.8	14	9	ADY64334	Ady64334	Thrombopo
608	45	80.4	14	4	AAU25986	Aau25986	Human thr	681	43	76.8	15	2	AAW66722	Aaw66722	Peptide c

682	43	76.8	15	2	AAW66720	Aaw66720 Peptide c	755	40	71.4	19	5	ABB72909	Abb72909 TPO mimet
683	43	76.8	15	2	AAW66728	Aaw66728 Peptide c	756	40	71.4	19	7	ADJ73061	Adj73061 TPO mimet
684	43	76.8	15	2	AAW66724	Aaw66724 Peptide c	757	40	71.4	19	8	ADJ52696	Adj52696 CH1 delet
685	43	76.8	15	4	AAU26033	Aau26033 Human thr	758	40	71.4	19	8	ADJ51657	Adj51657 CH1 delet
686	43	76.8	15	4	AAU25833	Aau25833 Human thr	759	40	71.4	329	4	ABG23127	Abg23127 Novel hum
687	43	76.8	15	4	AAU26027	Aau26027 Human thr	760	40	71.4	492	7	ABO82111	AbO82111 Pseudomon
688	43	76.8	15	4	AAU26025	Aau26025 Human thr	761	39	69.6	14	3	AAB17010	Aab17010 TPO-mimet
689	43	76.8	15	4	AAU26029	Aau26029 Human thr	762	39	69.6	14	5	ABB72896	Abb72896 TPO mimet
690	43	76.8	15	8	ADM72510	Adm72510 TPO mimet	763	39	69.6	14	7	ADJ73047	Adj73047 TPO mimet
691	43	76.8	15	8	ADM72506	Adm72506 TPO mimet	764	39	69.6	14	8	ADJ52682	Adj52682 CH1 delet
692	43	76.8	15	8	ADM72500	Adm72500 TPO mimet	765	39	69.6	14	8	ADJ51643	Adj51643 CH1 delet
693	43	76.8	15	8	ADM72518	Adm72518 TPO mimet	766	39	69.6	95	5	AAU09166	Aau09166 Human cyt
694	43	76.8	15	8	ADM72477	Adm72477 TPO mimet	767	39	69.6	100	5	ABP47800	Abp47800 HIV-2prot
695	43	76.8	15	8	ADM72504	Adm72504 TPO mimet	768	39	69.6	168	4	ABB65448	Abb65448 Drosophil
696	43	76.8	15	9	ADY64335	Ady64335 Thrombopo	769	39	69.6	230	2	AAR60767	Aar60767 HIV-2 ROD
697	43	76.8	16	2	AAW66711	Aaw66711 Peptide c	770	39	69.6	316	9	ABM93583	Abm93583 M. xanthu
698	43	76.8	16	8	ADM72476	Adm72476 TPO mimet	771	39	69.6	318	8	ADY22882	Ady22882 Plant ful
699	43	76.8	18	2	AAW09488	Aaw09488 Thrombopo	772	39	69.6	428	5	ABP47757	Abp47757 Protein #
700	43	76.8	18	2	AAW36639	Aaw36639 Thrombopo	773	39	69.6	450	3	AAG50584	Aag50584 Arabidops
701	43	76.8	18	4	AAU25858	Aau25858 Human thr	774	39	69.6	475	3	AAG50583	Aag50583 Arabidops
702	43	76.8	18	7	ADN59677	Adn59677 Thrombopo	775	39	69.6	499	3	AAG50582	Aag50582 Arabidops
703	43	76.8	25	7	ADN59734	Adn59734 Thrombopo	776	39	69.6	876	1	ABM92266	Abm92266 M. xanthu
704	43	76.8	28	8	ADJ52647	Adj52647 CH1 delet	777	39	69.6	1014	1	AAP80810	Aap80810 Sequence
705	43	76.8	29	3	AAB16972	Aab16972 TPO-mimet	778	39	69.6	1027	1	AAP81773	Aap81773 Sequence
706	43	76.8	29	5	ABB72858	Abb72858 TPO mimet	779	39	69.6	1035	2	AAR04025	Aar04025 Pol gene
707	43	76.8	29	7	ADJ73012	Adj73012 TPO mimet	780	39	69.6	1036	2	AAR20599	Aar20599 ROD HIV-2
708	43	76.8	29	7	ADJ73008	Adj73008 TPO mimet	781	39	69.6	1036	2	AAY51978	Aay51978 HIV-2 ROD
709	43	76.8	29	8	ADJ52643	Adj52643 CH1 delet	782	39	69.6	1036	6	ADA07961	Ada07961 HIV-2 pol
710	43	76.8	29	8	ADJ51604	Adj51604 CH1 delet	783	39	69.6	1036	8	ADL23699	Adl23699 HIV-2 ROD
711	43	76.8	29	8	ADJ51608	Adj51608 CH1 delet	784	38	67.9	10	2	AAW09473	Aaw09473 Thrombopo
712	43	76.8	30	9	ADY64336	Ady64336 Thrombopo	785	38	67.9	10	2	AAW36624	Aaw36624 Thrombopo
713	43	76.8	36	3	AAAY96521	Aay96521 Cyclic or	786	38	67.9	10	4	AAU25843	Aau25843 Human thr
714	43	76.8	36	3	AAB17298	Aab17298 TPO-mimet	787	38	67.9	28	7	AAE38642	Aae38642 Human col
715	43	76.8	36	3	AAB17299	Aab17299 TPO-mimet	788	38	67.9	311	4	AAM41798	Aam41798 Human pol
716	42	75.0	18	2	AAW09592	Aaw09592 Thrombopo	789	38	67.9	496	9	ABM91663	Abm91663 M. xanthu
717	42	75.0	18	2	AAW36743	Aaw36743 Thrombopo	790	38	67.9	549	2	AAW40113	Aaw40113 Human alp
718	42	75.0	18	4	AAU25962	Aau25962 Human thr	791	38	67.9	572	7	ABO80566	AbO80566 Pseudomon
719	42	75.0	36	3	AAAY96522	Aay96522 Linear th	792	38	67.9	694	8	ADQ67654	Adq67654 Novel hum
720	42	75.0	36	3	AAB17300	Aab17300 TPO-mimet	793	38	67.9	994	8	ABM82875	Abm82875 Human dia
721	42	75.0	100	5	ABP47801	Abp47801 Sooty man	794	38	67.9	1061	8	ABM82874	Abm82874 Human dia
722	42	75.0	429	5	ABP47758	Abp47758 Protein #	795	38	67.9	1078	4	AAM40012	Aam40012 Human pol
723	42	75.0	519	7	ADM05082	Adm05082 Human pro	796	38	67.9	1712	5	AAO17361	Aao17361 Human alp
724	41	73.2	18	7	ADN59679	Adn59679 Thrombopo	797	38	67.9	1712	8	ADS10473	Ads10473 Human the
725	41	73.2	18	7	ADN59676	Adn59676 Thrombopo	798	38	67.9	1712	8	ADU06526	Adu06526 Novel bro
726	41	73.2	22	7	ADN59662	Adn59662 Thrombopo	799	38	67.9	1712	9	ADY59885	Ady59885 Human col
727	41	73.2	25	7	ADN59829	Adn59829 TMP pepti	800	37.5	67.0	131	7	ABO67862	AbO67862 Pseudomon
728	41	73.2	25	7	ADN59738	Adn59738 Thrombopo	801	37.5	67.0	169	7	ABO76480	AbO76480 Pseudomon
729	41	73.2	25	7	ADN59706	Adn59706 Thrombopo	802	37	66.1	10	2	AAW35427	Aaw35427 Thrombopo
730	41	73.2	25	7	ADN59732	Adn59732 Thrombopo	803	37	66.1	10	4	AAU26002	Aau26002 Human thr
731	41	73.2	40	7	ADN59760	Adn59760 Peptide-v	804	37	66.1	14	2	AAW09479	Aaw09479 Thrombopo
732	41	73.2	302	6	ABU33828	Abu33828 Protein e	805	37	66.1	14	2	AAW09481	Aaw09481 Thrombopo
733	41	73.2	308	4	AAG91969	Aag91969 C glutami	806	37	66.1	14	2	AAW36630	Aaw36630 Thrombopo
734	40	71.4	14	2	AAW09484	Aaw09484 Thrombopo	807	37	66.1	14	2	AAW36632	Aaw36632 Thrombopo
735	40	71.4	14	2	AAW36766	Aaw36766 Thrombopo	808	37	66.1	14	2	AAW36631	Aaw36631 Thrombopo
736	40	71.4	14	2	AAW36635	Aaw36635 Thrombopo	809	37	66.1	14	2	AAW36631	Aaw36631 Thrombopo
737	40	71.4	14	3	AAB17009	Aab17009 TPO-mimet	810	37	66.1	14	3	AAB17013	Aab17013 TPO-mimet
738	40	71.4	14	4	AAU25854	Aau25854 Human thr	811	37	66.1	14	4	AAU25849	Aau25849 Human thr
739	40	71.4	14	5	ABB72895	Abb72895 TPO mimet	812	37	66.1	14	4	AAU25851	Aau25851 Human thr
740	40	71.4	14	7	ADJ73046	Adj73046 TPO mimet	813	37	66.1	14	5	ABB72899	Abb72899 TPO mimet
741	40	71.4	14	8	ADJ52681	Adj52681 CH1 delet	814	37	66.1	14	7	ADJ73050	Adj73050 TPO mimet
742	40	71.4	14	8	ADJ51642	Adj51642 CH1 delet	815	37	66.1	14	8	ADJ52685	Adj52685 CH1 delet
743	40	71.4	14	8	ADM72513	Adm72513 TPO mimet	816	37	66.1	14	8	ADJ51646	Adj51646 CH1 delet
744	40	71.4	15	2	AAW66725	Aaw66725 Peptide c	817	37	66.1	14	8	ADM72515	Adm72515 TPO mimet
745	40	71.4	15	4	AAU25834	Aau25834 Human thr	818	37	66.1	14	8	ADM72521	Adm72521 TPO mimet
746	40	71.4	15	4	AAU26030	Aau26030 Human thr	819	37	66.1	15	2	AAW66726	Aaw66726 Peptide c
747	40	71.4	15	8	ADM72512	Adm72512 TPO mimet	820	37	66.1	15	2	AAW66729	Aaw66729 Peptide c
748	40	71.4	18	2	AAW09490	Aaw09490 Thrombopo	821	37	66.1	15	4	AAU26031	Aau26031 Human thr
749	40	71.4	18	2	AAW36641	Aaw36641 Thrombopo	822	37	66.1	15	4	AAU26034	Aau26034 Human thr
750	40	71.4	18	4	AAU25860	Aau25860 Human thr	823	37	66.1	15	8	ADM72520	Adm72520 TPO mimet
751	40	71.4	19	2	AAW09495	Aaw09495 Thrombopo	824	37	66.1	15	8	ADM72514	Adm72514 TPO mimet
752	40	71.4	19	2	AAW36646	Aaw36646 Thrombopo	825	37	66.1	132	7	ADM04387	Adm04387 Human pro
753	40	71.4	19	3	AAB17023	Aab17023 TPO-mimet	826	37	66.1	137	8	ABO84792	AbO84792 Murine ca
754	40	71.4	19	4	AAU25865	Aau25865 Human thr	827	37	66.1	180	7	ABO80306	Pseudomon

828	37	66.1	261	7	ADH86920	Adh86920 Enterococ	901	36	64.3	513	9	AEB49395	Aeb49395 N. gonorr
829	37	66.1	306	5	ABG91631	Abg91631 Purine/py	902	36	64.3	513	9	AEB49397	Aeb49397 N. gonorr
830	37	66.1	306	5	ABG91624	Abg91624 Purine/py	903	36	64.3	520	7	ABO64867	Abo64867 Klebsiell
831	37	66.1	306	6	ABU36556	Abu36556 Protein e	904	36	64.3	592	4	ABG06780	Abg06780 Novel hum
832	37	66.1	306	6	ABU34427	Abu34427 Protein e	905	36	64.3	603	8	ADS25043	Ads25043 Bacterial
833	37	66.1	306	6	ABU35738	Abu35738 Protein e	906	36	64.3	607	3	AAy58580	Aay58580 Sorangium
834	37	66.1	321	8	ADX65923	Adx65923 Plant ful	907	36	64.3	685	6	ADA54712	Ada54712 Human pro
835	37	66.1	355	6	ABR58696	Abr58696 Human can	908	36	64.3	806	9	ABM93223	Abm93223 M. xanthu
836	37	66.1	355	7	ADF13714	Adf13714 Tumor-Ass	909	36	64.3	852	2	AAW79065	Aaw79065 FIV-NCSU1
837	37	66.1	355	8	ADS00100	Ads00100 Human ocu	910	36	64.3	854	2	AAW79065	Aaw79065 Feline Im
838	37	66.1	355	8	ADS00098	Ads00098 Human ocu	911	36	64.3	854	2	AAW79065	Aaw79065 Feline Im
839	37	66.1	355	9	ADW14773	Adw14773 Tumor-ass	912	36	64.3	854	2	AAW79065	Aaw79065 Feline Im
840	37	66.1	364	6	ABU23456	Abu23456 Protein e	913	36	64.3	855	2	AAR51250	Aar51250 FIV PPR e
841	37	66.1	372	4	AAG80999	Aag80999 HIV prote	914	36	64.3	855	2	AAR51250	Aar51250 Env polyp
842	37	66.1	476	8	ADT58335	Adt58335 Plant pol	915	36	64.3	855	2	AAR51252	Aar51252 FIV UK2 e
843	37	66.1	491	8	ADG76665	Adg76665 Phospholi	916	36	64.3	855	2	AAR51251	Aar51251 FIV JAPAN
844	37	66.1	526	8	ADY07792	Ady07792 Plant ful	917	36	64.3	855	2	AAR51255	Aar51255 FIV SWISS
845	37	66.1	777	8	ADS27768	Ads27768 Bacterial	918	36	64.3	855	2	AAR58585	Aar58585 FIV UK8 e
846	37	66.1	861	9	ADW38367	Adw38367 Ancestral	919	36	64.3	856	2	AAW23711	Aaw23711 Feline im
847	37	66.1	861	9	ADW38369	Adw38369 Ancestral	920	36	64.3	856	2	AAW23711	Aaw23711 Feline im
848	37	66.1	861	9	ADW38365	Adw38365 Ancestral	921	36	64.3	856	2	AAW23711	Aaw23711 Feline im
849	37	66.1	999	9	ADX40028	Adx40028 HIV Pol p	922	36	64.3	856	2	AAW23711	Aaw23711 Feline im
850	37	66.1	1055	3	AAAB12993	Aab12993 HIV-2 rev	923	36	64.3	856	2	AAW23711	Aaw23711 Feline im
851	36	64.3	14	2	AAW09478	Aaw09478 Thrombopo	924	36	64.3	856	2	AAW23711	Aaw23711 Feline im
852	36	64.3	14	2	AAW36629	Aaw36629 Thrombopo	925	36	64.3	856	2	AAW23711	Aaw23711 Feline im
853	36	64.3	14	4	AAU25848	Aau25848 Human thr	926	36	64.3	856	2	AAW23711	Aaw23711 Feline im
854	36	64.3	15	4	AAU26016	Aau26016 Human thr	927	36	64.3	856	2	AAW23711	Aaw23711 Feline im
855	36	64.3	15	8	ADM72480	Adm72480 TPO mimet	928	36	64.3	856	2	AAW23711	Aaw23711 Feline im
856	36	64.3	15	8	ADM72481	Adm72481 TPO mimet	929	36	64.3	856	2	AAW23711	Aaw23711 Feline im
857	36	64.3	15	9	ADV21762	Adv21762 SIV pol p	930	36	64.3	856	2	AAW23711	Aaw23711 Feline im
858	36	64.3	15	9	ADV21761	Adv21761 SIV pol p	931	36	64.3	856	2	AAW23711	Aaw23711 Feline im
859	36	64.3	16	4	AAU26017	Aau26017 Human thr	932	36	64.3	856	2	AAW23711	Aaw23711 Feline im
860	36	64.3	18	2	AAW09489	Aaw09489 Thrombopo	933	36	64.3	856	2	AAW23711	Aaw23711 Feline im
861	36	64.3	18	2	AAW36640	Aaw36640 Thrombopo	934	36	64.3	856	2	AAW23711	Aaw23711 Feline im
862	36	64.3	18	4	AAU25859	Aau25859 Human thr	935	36	64.3	856	2	AAW23711	Aaw23711 Feline im
863	36	64.3	110	4	AAW85655	Aam85655 Human imm	936	36	64.3	856	2	AAW23711	Aaw23711 Feline im
864	36	64.3	137	4	AAW91823	Aam91823 Human imm	937	36	64.3	856	2	AAW23711	Aaw23711 Feline im
865	36	64.3	142	8	ADX74465	Adx74465 Plant ful	938	36	64.3	856	2	AAW23711	Aaw23711 Feline im
866	36	64.3	169	7	ABM86829	Abm86829 Rice abio	939	36	64.3	856	2	AAW23711	Aaw23711 Feline im
867	36	64.3	176	5	AAE25747	Aae25747 Rice Ft h	940	36	64.3	856	2	AAW23711	Aaw23711 Feline im
868	36	64.3	176	8	ADR04263	Adr04263 Rice Ft h	941	36	64.3	856	2	AAW23711	Aaw23711 Feline im
869	36	64.3	200	8	ADN23470	Adn23470 Bacterial	942	36	64.3	856	2	AAW23711	Aaw23711 Feline im
870	36	64.3	254	8	ADX74695	Adx74695 Plant ful	943	36	64.3	856	2	AAW23711	Aaw23711 Feline im
871	36	64.3	268	6	ABM70732	Abm70732 Photorhab	944	36	64.3	856	2	AAW23711	Aaw23711 Feline im
872	36	64.3	271	8	ADS42097	Ads42097 Bacterial	945	36	64.3	856	2	AAW23711	Aaw23711 Feline im
873	36	64.3	307	3	AAG51074	Aag51074 Arabidops	946	36	64.3	856	2	AAW23711	Aaw23711 Feline im
874	36	64.3	307	3	AAG14587	Aag14587 Arabidops	947	36	64.3	856	2	AAW23711	Aaw23711 Feline im
875	36	64.3	312	5	ABB55368	Abb55368 Lactococc	948	36	64.3	856	2	AAW23711	Aaw23711 Feline im
876	36	64.3	316	3	AAG14586	Aag14586 Arabidops	949	36	64.3	856	2	AAW23711	Aaw23711 Feline im
877	36	64.3	316	3	AAG51073	Aag51073 Arabidops	950	36	64.3	856	2	AAW23711	Aaw23711 Feline im
878	36	64.3	316	5	ABP61149	Abp61149 Human GPC	951	36	64.3	856	2	AAW23711	Aaw23711 Feline im
879	36	64.3	316	7	ADC79356	Adc79356 Human G p	952	36	64.3	856	2	AAW23711	Aaw23711 Feline im
880	36	64.3	324	4	AAU24556	Aau24556 Human olf	953	36	64.3	856	2	AAW23711	Aaw23711 Feline im
881	36	64.3	324	5	ABP61135	Abp61135 Human GPC	954	36	64.3	856	2	AAW23711	Aaw23711 Feline im
882	36	64.3	324	5	ABP95673	Abp95673 Human olf	955	36	64.3	856	2	AAW23711	Aaw23711 Feline im
883	36	64.3	324	5	AAU95749	Aau95749 Human olf	956	36	64.3	856	2	AAW23711	Aaw23711 Feline im
884	36	64.3	324	5	AAU85176	Aau85176 G-coupled	957	36	64.3	856	2	AAW23711	Aaw23711 Feline im
885	36	64.3	324	7	ADC86379	Adc86379 Human GPC	958	36	64.3	856	2	AAW23711	Aaw23711 Feline im
886	36	64.3	324	8	ADG83490	Adg83490 Human Olf	959	36	64.3	856	2	AAW23711	Aaw23711 Feline im
887	36	64.3	324	8	ADG79328	Adg79328 Human G p	960	36	64.3	856	2	AAW23711	Aaw23711 Feline im
888	36	64.3	325	7	ADC51072	Aag51072 Arabidops	961	36	64.3	856	2	AAW23711	Aaw23711 Feline im
889	36	64.3	331	3	AAG14585	Aag14585 Arabidops	962	36	64.3	856	2	AAW23711	Aaw23711 Feline im
890	36	64.3	350	2	AAW62621	Aaw62621 Mus muscu	963	36	64.3	856	2	AAW23711	Aaw23711 Feline im
891	36	64.3	371	8	ADM98610	Adm98610 Geranylge	964	36	64.3	856	2	AAW23711	Aaw23711 Feline im
892	36	64.3	408	5	AAE16177	Aae16177 Human G-p	965	36	64.3	856	2	AAW23711	Aaw23711 Feline im
893	36	64.3	486	6	ABP77210	Abp77210 N. gonorr	966	36	64.3	856	2	AAW23711	Aaw23711 Feline im
894	36	64.3	512	4	ABG25367	Abg25367 Novel hum	967	36	64.3	856	2	AAW23711	Aaw23711 Feline im
895	36	64.3	513	2	AAy38789	Aay38789 Neisseria	968	36	64.3	856	2	AAW23711	Aaw23711 Feline im
896	36	64.3	513	2	AAy38786	Aay38786 Neisseria	969	36	64.3	856	2	AAW23711	Aaw23711 Feline im
897	36	64.3	513	2	AAy38787	Aay38787 Neisseria	970	36	64.3	856	2	AAW23711	Aaw23711 Feline im
898	36	64.3	513	2	AAy38788	Aay38788 Neisseria	971	36	64.3	856	2	AAW23711	Aaw23711 Feline im
899	36	64.3	513	9	AEB49392	Aeb49392 N. mening	972	36	64.3	856	2	AAW23711	Aaw23711 Feline im
900	36	64.3	513	9	AEB49394	Aeb49394 N. mening	973	36	64.3	856	2	AAW23711	Aaw23711 Feline im

974 35 62.5 87 4 AAU63256
975 35 62.5 87 6 ABMS9775
976 35 62.5 120 4 ABG07816
977 35 62.5 212 7 ABO68391
978 35 62.5 235 3 AAG57785
979 35 62.5 235 3 AAG57786
980 35 62.5 235 3 AAG55516
981 35 62.5 235 3 AAG57791
982 35 62.5 241 3 AAG17852
983 35 62.5 244 6 ABU23180
984 35 62.5 250 3 AAG17851
985 35 62.5 302 3 AAG55515
986 35 62.5 303 2 AAR99488
987 35 62.5 306 4 ABG22498
988 35 62.5 310 9 ABM96497
989 35 62.5 315 6 ABU21946
990 35 62.5 318 4 AAU19231
991 35 62.5 318 4 AAU19230
992 35 62.5 325 7 ADD24938
993 35 62.5 326 3 AAG17850
994 35 62.5 353 9 ABM95910
995 35 62.5 391 3 AAG30100
996 35 62.5 408 3 AAG57790
997 35 62.5 413 3 AAG57789
998 35 62.5 422 3 AAY96815
999 35 62.5 433 8 ADT87825
1000 35 62.5 455 3 AAG59392

Aau63256 Propionib
Abms9775 Propionib
Abg07816 Novel hum
Abo68391 Pseudomon
Aag57785 Arabidops
Aag57786 Arabidops
Aag55516 Arabidops
Aag57791 Arabidops
Aag17852 Arabidops
Abu23180 Protein e
Aag17851 Arabidops
Aag55515 Arabidops
Aar99488 PcpR. 10/
Abg22498 Novel hum
Abm96497 M. xanthu
Abu21946 Protein e
Aau19231 Human G p
Aau19230 Human G p
Add24938 Escherich
Aag17850 Arabidops
Abm95910 M. xanthu
Aag30100 Arabidops
Aag57790 Arabidops
Aag57789 Arabidops
Aay96815 A. thalia
Adt87825 Plant hom
Aags9392 Arabidops

ALIGNMENTS

RESULT 1
AAB17006
ID AAB17006 standard; peptide; 10 AA.
XX
AC AAB17006;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:62.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS
XX WO200024782-A2.
PN
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 216; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P³, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention

SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||||||
1 EGPTLRQWLA 10

RESULT 2
ABB72892
ID ABB72892 standard; peptide; 10 AA.
XX
AC ABB72892;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:62.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW anastemic; anorectic; antiinfectivity; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.

PS Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 5; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

|||||

RESULT 3

ADJ73043

ID ADJ73043 standard; peptide; 10 AA.

XX

AC ADJ73043;

XX

DT 06-MAY-2004 (first entry)

DE TPO mimetic peptide sequence SeqID 497.

XX

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;

KW cardiovascular; infectious; malignant; neurologic disease; anaemia;

KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

KW TPO.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

PI

XX WPI; 2003-804237/75.

DR

XX New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX

.PS Disclosure; SEQ ID NO 497; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted, into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is a TPO mimetic peptide sequence used to make a

CC mimetibody of the invention.

XX

SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 7; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

|||||

RESULT 4

ADJ52678

ID ADJ52678 standard; peptide; 10 AA.

XX

AC ADJ52678;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID497.

XX

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PA (CENZ) CENTOCOR INC.

XX

XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082870/08.

XX

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX

PS Claim 2; SEQ ID NO 497; 129pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 1 EGPTLRQWLA 10
|||||

RESULT 5
ADJ51639
ID ADJ51639 standard; peptide; 10 AA.
XX
AC ADJ51639;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID497.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 497; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 1 EGPTLRQWLA 10
|||||

RESULT 6
AAW36787
ID AAW36787 standard; peptide; 12 AA.
XX
AC AAW36787;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cross-links 12
FT /note= "terminal carboxy group linked to epsilon amino
FT group of Lys14 in AAW36788"

XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US0009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX

PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 78; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 56; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
| | | | | | | |
Db 1 EGPTLRQWLA 10

RESULT 7
AAW36781
ID AAW36781 standard; peptide; 12 AA.
XX
AC AAW36781;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cross-links 12
FT /note= "terminal carboxy group linked to epsilon amino
FT group of Lys14 in AAW36782"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 78; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 56; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
| | | | | | | |
Db 2 EGPTLRQWLA 11

RESULT 8
AAU26014
ID AAU26014 standard; peptide; 12 AA.
XX
AC AAU26014;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #200.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 151; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living

cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

```
Query Match      100.0%; Score 56; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 1 EGPTLRQWLA 10

RESULT 9
AAW36792
ID AAW36792 standard: peptide: 13 AA.

Thrombopoietin receptor; binding peptide; treatment; agonist;
KW
haematological disorder; thrombocytopaenia; chemotherapy;
KW
radiation therapy; bone marrow transfusion; diagnosis;
KW
signal transduction; receptor activation; cell culture.
KW

Key	Location/Qualifiers
Modified-site	1
FT	
FT	/note= "optionally acylated"
FT	12
Modified-site	
FT	/label= bAla
FT	13
Cross-links	
FT	/note= "epsilon amino group of Lys13 linked to terminal
FT	carboxy group of AAW36791"
FT	

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

Peptides and peptide mimetics which bind to and activate the thrombopoietin receptor - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

The present peptide, which binds the thrombopoietin receptor (TR), can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopaenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 56; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 1 EGPTLRQWLA 10

RESULT 10
AAW36779
ID AAW36779 standard; peptide; 13 AA.

KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

	Key	Location/Qualifiers
FT	FH	13
FT	Cross-links	/note="terminal carboxy group linked to epsilon amino group of Lys15 in AAW36780"
FT		
FT		

....
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

Peptides and peptide mimetics which bind to and activate the thrombopoietin receptor - useful in treatment of haematological disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

The present peptide, which binds the thrombopoietin receptor (TR), can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopaenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

Query Match 100.0%; Score 56; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Qy 1 EGPTLRQWLA 10

Db		2	EGPTLRQWLA	11
RESULT 11				
AAU26008				
ID	AAU26008	standard; peptide; 13	AA.	
XX				
AC	AAU26008;			
XX				
DT	17-DEC-2001	(first entry)		
XX				
DE	Human thrombopoietin receptor (TPO-R)	activator peptide #194.		
XX				
KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;			
KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;			
KW	bone marrow transplantation; haematological disorder; platelet disorder;			
KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;			
KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;			
KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.			
XX				
OS	Homo sapiens.			
XX				
PN	US6251864-B1.			
XX				
PD	26-JUN-2001.			
XX				
PF	01-MAR-2000; 2000US-00516704.			
XX				
PR	07-JUN-1995; 95US-00478128.			
PR	07-JUN-1995; 95US-00485301.			
PR	07-JUN-1996; 96WO-US009623.			
PR	15-AUG-1996; 96US-00699027.			
XX				
PA	(GLAX) GLAXO GROUP LTD.			
XX				
PI	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;			
PI	Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;			
PI	Yin Q;			
XX				
DR	WPI; 2001-564142/63.			
XX				
PT	Activating thrombopoietin receptors in cells, used to treat			
PT	thrombocytopenia and hematological disorders, comprises contacting cells			
PT	with peptides and peptide mimetics attached to hydrophilic polymers.			
XX				
PS	Disclosure; Col 147; 128pp; English.			
XX				
CC	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that			
CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods			
CC	of activating thrombopoietin receptors in cells comprise contacting the			
CC	cells with effective amounts of peptides and peptide mimetics attached to			
CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such			
CC	as that due to chemotherapy, radiation therapy or bone-marrow			
CC	transplantation and to prevent thrombocytopenia in patients at risk.The			
CC	sequences are used to treat and prevent haematological disorders			
CC	including thrombocytopenia and platelet disorders. They are used in vitro			
CC	as unique tools for understanding the biological role of thrombopoietin			
CC	(TPO) and to develop other compounds that bind to and activate the TPO			
CC	receptor. The peptides can be used to detect TPO receptors on living			
CC	cells and fixed cells, in biological fluids, in tissue homogenates, and			
CC	in purified or natural biological materials. They may also be used for in			
CC	situ staining, fluorescence-activated cell sorting, Western blotting and			
CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can			
CC	be used for in vitro expansion of megakaryocytes and their committed			
CC	progenitors alone or in conjunction with additional cytokines			
XX				
SQ	Sequence 13 AA;			
Query Match	100.0%;	Score 56;	DB 4;	Length 13;
Best Local Similarity	100.0%;	Pred. No. 0.0055;		
Matches	10;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

OY	1	EGPTLRQWLA	10
Db	2	EGPTLRQWLA	11
RESULT 12			
AAU26012			
ID	AAU26012	standard; peptide; 13	AA.
XX			
AC	AAU26012;		
XX			
DT	17-DEC-2001	(first entry)	
XX			
DE	Human thrombopoietin receptor (TPO-R)	activator peptide #198.	
XX			
KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;		
KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;		
KW	bone marrow transplantation; haematological disorder; platelet disorder;		
KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;		
KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;		
KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.		
XX			
OS	Homo sapiens.		
XX			
PN	US6251864-B1.		
XX			
PD	26-JUN-2001.		
XX			
PF	01-MAR-2000; 2000US-00516704.		
XX			
PR	07-JUN-1995; 95US-00478128.		
PR	07-JUN-1995; 95US-00485301.		
PR	07-JUN-1996; 96WO-US009623.		
PR	15-AUG-1996; 96US-00699027.		
XX			
PA	(GLAX) GLAXO GROUP LTD.		
XX			
PI	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;		
PI	Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;		
PI	Yin Q;		
XX			
DR	WPI; 2001-564142/63.		
XX			
PT	Activating thrombopoietin receptors in cells, used to treat		
PT	thrombocytopenia and hematological disorders, comprises contacting cells		
PT	with peptides and peptide mimetics attached to hydrophilic polymers.		
XX			
PS	Disclosure; Col 149-150; 128pp; English.		
XX			
CC	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that		
CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods		
CC	of activating thrombopoietin receptors in cells comprise contacting the		
CC	cells with effective amounts of peptides and peptide mimetics attached to		
CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such		
CC	as that due to chemotherapy, radiation therapy or bone-marrow		
CC	transplantation and to prevent thrombocytopenia in patients at risk.The		
CC	sequences are used to treat and prevent haematological disorders		
CC	including thrombocytopenia and platelet disorders. They are used in vitro		
CC	as unique tools for understanding the biological role of thrombopoietin		
CC	(TPO) and to develop other compounds that bind to and activate the TPO		
CC	receptor. The peptides can be used to detect TPO receptors on living		
CC	cells and fixed cells, in biological fluids, in tissue homogenates, and		
CC	in purified or natural biological materials. They may also be used for in		
CC	situ staining, fluorescence-activated cell sorting, Western blotting and		
CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can		
CC	be used for in vitro expansion of megakaryocytes and their committed		
CC	progenitors alone or in conjunction with additional cytokines		
XX			
SQ	Sequence 13 AA;		
Query Match	100.0%;	Score 56;	DB 4;
Best Local Similarity	100.0%;	Pred. No. 0.0055;	Length 13;
Matches	10;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	EGPTLRQWLA 10 	Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	1	EGPTLRQWLA 10					
RESULT 13							
AAU26015	AAU26015 standard; peptide; 13 AA.						
XX	AC	AAU26015;					
XX	DT	18-DEC-2001 (first entry)					
XX	DE	Human thrombopoietin receptor (TPO-R) activator peptide #201.					
XX	KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;					
KW	KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;					
KW	KW	bone marrow transplantation; haematological disorder; platelet disorder;					
KW	KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;					
KW	KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;					
KW	KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.					
XX	OS	Homo sapiens.					
XX	XX	US6251864-B1.					
PN	XX	26-JUN-2001.					
PD	XX						
XX	PF	01-MAR-2000; 2000US-00516704.					
XX	PR	07-JUN-1995; 95US-00478128.					
PR	PR	07-JUN-1995; 95US-00485301.					
PR	PR	07-JUN-1996; 96WO-US009623.					
PR	PR	15-AUG-1996; 96US-00699027.					
XX	PA	(GLAX) GLAXO GROUP LTD.					
XX	PI	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;					
PI	PI	Balesubramanian P, Wagstrom CR, Hendrien RW, Deprince RB, Podduturi S;					
PI	PI	Yin Q;					
XX	DR	WPI; 2001-564142/63.					
XX	PT	Activating thrombopoietin receptors in cells, used to treat					
PT	PT	thrombocytopenia and hematological disorders, comprises contacting cells					
PT	PT	with peptides and peptide mimetics attached to hydrophilic polymers.					
XX	PS	Disclosure; Col 151; 128pp; English.					
XX	CC	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that					
CC	CC	bind to and activate the human thrombopoietin receptor (TPO-R). Methods					
CC	CC	of activating thrombopoietin receptors in cells comprise contacting the					
CC	CC	cells with effective amounts of peptides and peptide mimetics attached to					
CC	CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such					
CC	CC	as that due to chemotherapy, radiation therapy or bone-marrow					
CC	CC	transplantation and to prevent thrombocytopenia in patients at risk.The					
CC	CC	sequences are used to treat and prevent haematological disorders					
CC	CC	including thrombocytopenia and platelet disorders. They are used in vitro					
CC	CC	as unique tools for understanding the biological role of thrombopoietin					
CC	CC	(TPO) and to develop other compounds that bind to and activate the TPO					
CC	CC	receptor. The peptides can be used to detect TPO receptors on living					
CC	CC	cells and fixed cells, in biological fluids, in tissue homogenates, and					
CC	CC	in purified or natural biological materials. They may also be used for in					
CC	CC	situ staining, fluorescence-activated cell sorting, Western blotting and					
CC	CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can					
CC	CC	be used for in vitro expansion of megakaryocytes and their committed					
CC	CC	progenitors alone or in conjunction with additional cytokines					
XX	XX						
SQ	SQ	Sequence 13 AA;					
Query Match 100.0%; Score 56; DB 4; Length 13;							
Best Local Similarity 100.0%; Pred. No. 0.0055;							
QY	1	EGPTLRQWLA 10 	Query Match 100.0%;	Score 56;	DB 7;	Length 13;	
Db	1	EGPTLRQWLA 10	Best Local Similarity 100.0%;	Pred. No. 0.0055;			
			Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
RESULT 14							
ADJ73003	ADJ73003 standard; peptide; 13 AA.						
XX	AC	ADJ73003;					
XX	DT	06-MAY-2004 (first entry)					
XX	DE	TPO mimetic peptide sequence SeqID 457.					
XX	KW	mimetic; CDR mimetibody; gene therapy; transgenic; immune;					
KW	KW	cardiovascular; infectious; malignant; neurologic disease; anaemia;					
KW	KW	immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;					
KW	KW	TPO.					
XX	OS	Synthetic.					
XX	XX	WO2003084477-A2.					
PN	XX	16-OCT-2003.					
PD	XX						
XX	PF	24-MAR-2003; 2003WO-US009139.					
XX	PR	29-MAR-2002; 2002US-0368791P.					
XX	PA	(CENZ) CENTOCOR INC.					
XX	PI	Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;					
XX	DR	WPI; 2003-804237/75.					
XX	PT	New CDR mimetibody comprising a portion of a heavy or light chain					
PT	PT	variable region comprising human framework or ligand binding region,					
PT	PT	useful for preparing a composition for treating e.g., immune,					
PT	PT	cardiovascular or neurologic disease.					
XX	PS	Disclosure; SEQ ID NO 457; 97pp; English.					
XX	CC	This invention relates to novel mammalian CDR mimetibodies, specific					
CC	CC	portions or variants thereof. Specifically, it refers to an antibody					
CC	CC	fragment where a protein has been inserted into, or replaces a portion					
CC	CC	of, one or more CDR regions, such that each CDR mimetibody comprises at					
CC	CC	least one portion of a heavy chain or light chain variable region, which					
CC	CC	itself comprises at least one human framework region and at least one					
CC	CC	ligand binding region (LBR). The present invention describes human					
CC	CC	mimetibodies, including modified immunoglobulins and cleavage products					
CC	CC	that can be useful in gene therapy and the generation of transgenic					
CC	CC	plants and animals. Furthermore, the CDR mimetibody is useful for					
CC	CC	preparing compositions for modulating, treating or reducing the symptoms					
CC	CC	of immune, cardiovascular, infectious, malignant and/ or neurologic					
CC	CC	diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,					
CC	CC	cardiant, antimicrobial, cytostatic and neuroprotective activities. This					
CC	CC	peptide sequence is a TPO mimetic peptide sequence used to make a					
CC	CC	mimetibody of the invention.					
XX	XX						
SQ	SQ	Sequence 13 AA;					
Query Match 100.0%; Score 56; DB 7; Length 13;							
Best Local Similarity 100.0%; Pred. No. 0.0055;							
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
QY	1	EGPTLRQWLA 10 					
Db	1	EGPTLRQWLA 10					

RESULT 15
ADJ52638
ID ADJ52638 standard; peptide; 13 AA.
XX
AC ADJ52638;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID457.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
PF 28-JUN-2002; 2002US-0392431P.
XX
PR (CENZ) CENTOCOR INC.
XX
PA Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX WPI; 2004-082870/08.
DR
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 457; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 56; DB 8; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
| | | | | | | |
Db 1 EGPTLRQWLA 10
RESULT 16
ADJ51599
ID ADJ51599 standard; peptide; 13 AA.

XX
AC ADJ51599;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID457.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
PF 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 14; SEQ ID NO 457; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 56; DB 8; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 1 EGPTLRQWLA 10

RESULT 17
AAW09463
ID AAW09463 standard; protein; 14 AA.
XX
AC AAW09463;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding compound peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..14
FT /note= "Preferably linkages are selected from: -
FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
FT lower alkyl"
FT Modified-site 1
FT /note= "Preferably N-terminus is selected from: -NRR1; -
FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
FT benzylloxycarbonyl-NH; benzylloxycarbonyl-NH with 1-3
FT substitutions on the phenyl ring selected from lower
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
FT independently selected from hydrogen and lower alkyl"
FT Modified-site 14
FT /note= "Preferably C-terminus is -C(O)R2 where R2 is
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
FT and R4 are independently selected from hydrogen and lower
FT alkyl, and where the nitrogen atom of the -NR3R4 group
FT can optionally be the amine group of the N-terminus of
FT the peptide forming a cyclic peptide"

XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PI Thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 18; Page 89; 106pp; English.
XX
CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding
CC affinity to TR as expressed by an IC50 of no more than about 100 mum. The
CC compound (especially if modified, see features table) can be used for
CC treating patients suffering from haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 2 EGPTLRQWLA 11

RESULT 18
AAW09468
ID AAW09468 standard; protein; 14 AA.
XX
AC AAW09468;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding compound peptide (part of a dimer).
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cross-links 14
FT /note= "Linked to the omega Lys from AAW19534"
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 30; Page 91; 106pp; English.
XX
CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It is part of a dimer linked by the omega amino acid to
CC the omega amino acid in the sequence in AAW19534. The compound can be
CC used for treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.006; Mismatches 0; Indels 0; Gaps 0; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11

RESULT 19
AAW33030
ID AAW33030 standard; peptide; 14 AA.
XX
AC AAW33030;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX WO9640750-A1.
PN
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 19; Page 89; 106pp; English.
XX
CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopaenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11

RESULT 20
AAW33034
ID AAW33034 standard; peptide; 14 AA.
XX

AAW33034;
11-MAR-1998 (first entry)
Thrombopoietin receptor binding peptide.
Thrombopoietin receptor; binding peptide; treatment; agonist;
haematological disorder; thrombocytopaenia; chemotherapy;
radiation therapy; bone marrow transfusion; diagnosis;
signal transduction; receptor activation; cell culture.
Synthetic.
Key Location/Qualifiers
Cross-links 14
/note= "terminal carboxy group linked to epsilon amino
group of Lys16 in AAW33035"
WO9640750-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US009623.
07-JUN-1995; 95US-00478128.
07-JUN-1995; 95US-00485301.
(GLAX) GLAXO GROUP LTD.
Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
WPI; 1997-052226/05.
Peptides and peptide mimetics which bind to and activate the
thrombopoietin receptor - useful in treatment of haematological
disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
Claim 30; Page 91; 106pp; English.
The present peptide binds the thrombopoietin receptor (TR), has a
molecular weight of less than 8000 Da and a TR binding affinity as
expressed by an IC50 of no more than about 100 microm. It can be used to
treat disorders which are susceptible to treatment with a thrombopoietin
agonist, preferably haematological disorders and thrombocytopaenia
resulting from chemotherapy, radiation therapy or bone marrow
transfusions. It can also be used diagnostically, e.g. to investigate the
mechanism of thrombopoietin signal transduction and receptor activation,
or to maintain the proliferation and growth of thrombopoietin dependent
cell lines
Sequence 14 AA;
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11

RESULT 21
AAW36782
ID AAW36782 standard; peptide; 14 AA.
XX
AC AAW36782;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 13
FT /label= bAla
FT Cross-links 14
FT /note= "epsilon amino group of Lys14 linked to terminal
FT carboxy group of AAW36781"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 78; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;
XX
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11

RESULT 22
AAW36774
ID AAW36774 standard; peptide; 14 AA.
XX
AC AAW36774;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

FT Modified-site 14
FT /note= "NH2-Ala"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 77; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;
XX
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11

RESULT 23
AAW36788
ID AAW36788 standard; peptide; 14 AA.
XX
AC AAW36788;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 13
FT /label= bAla
FT Cross-links 14
FT /note= "epsilon amino group of Lys14 linked to terminal
FT carboxy group of AAW36787"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX

PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
PI
XX WPI; 1997-0522226/05.
DR
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
PT
XX Example 9; Page 78; 106pp; English.
PS
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||
1 EGPTLRQWLA 10

RESULT 24
ADI24843
ID ADI24843 standard; peptide; 14 AA.
XX
AC ADI24843;
XX
DT 15-APR-2004 (first entry)
XX
DE AF 12505 as active moiety for pharmacologically active peptide.
XX
KW pharmacologically active peptide conjugate; enzymatic cleavage; pain;
KW HIV; cancer; diabetes; incontinence; hypertension; amnesia;
KW Alzheimer's disease; fever; depression; sex hormone regulation;
KW eating disorder; schizophrenia; osteoporosis; insomnia;
KW Central nervous system disorder; contraceptive.
XX
OS Synthetic.
XX
PN WO9946283-A1.
XX
XX 16-SEP-1999.
XX
PF 09-MAR-1999; 99WO-DK000118.
XX
PR 09-MAR-1998; 98DK-00000317.
XX
PA (ZEAL-) ZEALAND PHARM AS.
XX
XX Larsen BD;
PI
XX WPI; 1999-561659/47.
DR
XX New peptide conjugates used for treating, e.g. pain, HIV, depression,
PT schizophrenia, osteoporosis or insomnia.
PT
XX

PS Claim 24; Page 90; 113pp; English.
XX
CC The invention relates to a novel pharmacologically active peptide
CC conjugate having a reduced tendency towards enzymatic cleavage comprises
CC X and Z, where: (a) X is a pharmacologically active peptide sequence; and
CC (b) Z is a stabilising peptide sequence of 4-20 amino acid units
CC covalently bound to X, where each amino acid unit in the stabilizing
CC peptide sequence, Z being selected from Ala, Leu, Ser, Thr, Tyr, Asn,
CC Gln, Asp, Glu, Lys, Arg, His, Met, Orn, and amino acid units of formula -
CC NH-C(R1)(R2)-C(=O)- (I), where: R1 and R2 are H, 1-6C alkyl, phenyl, and
CC phenyl-methyl, where 1-6C-alkyl is optionally substituted with 1 -3
CC substituents selected from halogen, hydroxy, amino, cyano, nitro,
CC sulfonyl, and carboxy, and phenyl and phenyl-methyl are optionally
CC substituted with 1-3 substituents selected from 1-6C-alkyl, 2-6C-alkenyl,
CC halogen, hydroxy, amino, cyano, nitro, sulfonyl, and carboxy, or R 1 and R
CC 2 together with the C atom to which they are bound form a cyclopentyl,
CC cyclohexyl or cycloheptyl ring, e.g. 2,4-diaminobutanoic acid and 2,3-
CC diaminopropanoic acid; the ratio between the half-life of the peptide
CC conjugate and the half-life of the corresponding active peptide sequence,
CC X, when treated with carboxypeptidase A or leucine aminopeptidase in
CC about 50 mM phosphate buffer solution at about pH 7.4 and about 37 deg C
CC or in serum or plasma is at least about 2 (preferably at least about 10),
CC or when the pharmacologically active peptide X is not orally absorbed,
CC the conjugate is adsorbed, or a salt, with the proviso that the
CC pharmacologically active peptide conjugate is not selected from sequences
CC (ADI24837)-(ADI24841). The peptide conjugates can be used for treating
CC e.g. pain, HIV, cancer, diabetes, incontinence, hypertension, amnesia,
CC Alzheimer's disease, fever, depression, sex hormone regulation, eating
CC disorders, schizophrenia, osteoporosis or insomnia. They can also be used
CC for treating e.g. CNS disorders and as contraceptives. The conjugated
CC peptides are less susceptible to degradation by proteases compared to the
CC corresponding free pharmacologically active peptides. This sequence
CC represents a pharmacologically active peptide as the X part of the
CC peptide of the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11

RESULT 25
AA96515
ID AA96515 standard; peptide; 14 AA.
XX
AC AA96515;
XX
DT 04-SEP-2000 (first entry)
XX
DE Thrombopoietin mimetic peptide.
XX
KW Thrombopoietin; mimetic; TMP; TP0; platelet; megakaryocyte; production;
KW anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;
KW immunosuppressive; anti-inflammatory.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 14
FT /note= "subunits in the dimer are covalently bonded at
FT each carboxy terminus through peptide linkage with NH2-
FT (CH2)4-CH(CONH2)-NH-CO-(CH2)2-NH2"
XX
PN WO200024770-A2.
XX
PD 04-MAY-2000.
XX
PF 22-OCT-1999; 99WO-US024834.

XX 23-OCT-1998; 98US-0105348P.
XX (AMGE-) AMGEN INC.
PA Liu C, Feige U, Cheetham J;
XX WPI; 2000-365108/31.
XX Thrombopoietic peptides which activate mpl receptors and increase the
PT production of platelets or platelet precursors, useful for treatment of
PT diseases which involve thrombocytopenia.
XX Claim 7; Page 60; 91pp; English.
XX A compound which binds to an mpl receptor comprising a thrombopoietin
CC mimetic peptide (TMP) dimer joined by a linker [TMP 1-(L 1) nTMP 2], is
CC new. TMP 1 and TMP 2 are amino acid sequences varying from at least 10 to
CC 14 residues in length comprising X 2-X 1 0, X 2-X 1 1, X 2-X 1 2, X 2-
CC X 1 3, X 2-X 1 4, X 1-X 1 0, X 1-X 1 1, X 1-X 1 2, X 1-X 1 3, and X 1-
CC X 1 4. X 1 = I, A, V, L, S or R; X 2 = E, D, K or V; X 3 = G or A; X 4 =
CC P; X 5 = T or S; X 6 = L, I, V, A or F; X 7 = R or K; X 8 = Q, N, or E;
CC X 9 = W, Y or F; X 1 0 = L, I, V, A, F, M, or K; X 1 1 = A, I, V, L, F,
CC S, T, K, H, or E; X 1 2 = A, I, V, L, F, G, S, or Q; X 1 3 = R, K, T, V,
CC N, Q or G; X 1 4 = A, I, V, L, F, T, R, E, or G; L 1 = linker comprising
CC 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate
CC the c-Mpl receptor which mediates the activity of endogenous
CC thrombopoietin. The TMPs are useful for increasing the production of
CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which
CC is useful for treatment of diseases which involve thrombocytopenia, e.g.
CC aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency
CC virus associated ITP, and systemic lupus erythematosus
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11
RESULT 26
AAB16962
ID AAB16962 standard; peptide; 14 AA.
XX
AC AAB16962;
XX 31-OCT-2000 (first entry)
XX TPO-mimetic peptide TMP SEQ ID NO:13.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS WO200024782-A2.
XX 04-MAY-2000.
XX 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.

PA (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX Claim 19; Page 189; 608pp; English.
PS The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX Sequence 14 AA;
SQ Query Match 100.0%; Score 56; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11
RESULT 27
AAB16968
ID AAB16968 standard; peptide; 14 AA.
XX
AC AAB16968;
XX 31-OCT-2000 (first entry)
XX TPO-mimetic peptide sequence SEQ ID NO:24.
DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX Synthetic.
OS WO200024782-A2.
XX 04-MAY-2000.
XX 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.
XX (AMGE-) AMGEN INC.
PA Feige U, Liu C, Cheetham J, Boone TC;
XX

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 203; 608pp; English.

XX The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
|||||
Db 2 EGPTLRQWLA 11

RESULT 28
AAU26009
ID AAU26009 standard; peptide; 14 AA.

XX AAU26009;

XX 18-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #195.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine; haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA; bone marrow transplantation; haematological disorder; platelet disorder; enzyme-linked immunosorbent assay; in situ staining; biological fluid; tissue homogenate; fluorescence-activated cell sorting; Western blotting; in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S; Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat thrombocytopenia and hematological disorders, comprises contacting cells with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 149-150; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin (TPO) and to develop other compounds that bind to and activate the TPO receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
|||||
Db 2 EGPTLRQWLA 11

RESULT 29
AAU26013
ID AAU26013 standard; peptide; 14 AA.

XX AAU26013;

XX 18-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #199.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine; haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA; bone marrow transplantation; haematological disorder; platelet disorder; enzyme-linked immunosorbent assay; in situ staining; biological fluid; tissue homogenate; fluorescence-activated cell sorting; Western blotting; in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI Yin Q;
XX WPI; 2001-564142/63.
DR
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 151-152; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
|||
Db 1 EGPTLRQWLA 10

RESULT 30
AAU26006
ID AAU26006 standard; peptide; 14 AA.

XX AC AAU26006;
XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #192.
DE
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
DR
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 147; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
|||
Db 2 EGPTLRQWLA 11

RESULT 31
AAU25827
ID AAU25827 standard; peptide; 14 AA.

XX AC AAU25827;
XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #13.
DE
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX

PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
DR WPI; 2001-564142/63.
XX
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 69-70; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
|||
Db 2 EGPTLRQWLA 11

RESULT 32
AAU26010
ID AAU26010 standard; peptide; 14 AA.

XX
AC AAU26010;

XX
DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #196.

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
DR
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 149-150; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
|||
Db 1 EGPTLRQWLA 10

RESULT 33
AAU26004
ID AAU26004 standard; peptide; 14 AA.

XX
AC AAU26004;

XX
DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #190.

KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX

PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 147; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent hematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
| | | | | | | | | |
| | | | | | | | | |

RESULT 34
ABB72854
ID ABB72854 standard; peptide; 14 AA.
XX
AC ABB72854;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:24.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.
PF
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
| | | | | | | | | |
| | | | | | | | | |

RESULT 35
ABB72853
ID ABB72853 standard; peptide; 14 AA.
XX
AC ABB72853;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:13.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS Synthetic.
OS WO200183525-A2.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-US014310.
XX PR 03-MAY-2000; 2000US-005633286.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
DR Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
PT
XX
PS Claim 39; Page 43; 176pp; English.
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11
RESULT 36
ABP51669
ID ABP51669 standard; peptide; 14 AA.
XX
AC ABP51669;
XX
DT 01-OCT-2002 (first entry)
XX
DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:1.
XX
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX
OS Homo sapiens.

OS Synthetic.
XX WO200246238-A2.
XX PD 13-JUN-2002.
XX PF 05-DEC-2001; 2001WO-US047656.
XX PR 05-DEC-2000; 2000US-0251448P.
XX PR 04-MAY-2001; 2001US-0288889P.
XX PR 29-MAY-2001; 2001US-0294068P.
XX PA (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX WPI; 2002-566610/60.
DR A novel immunogen molecule comprising a region in which amino acid
XX residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoietin mimetic.
XX
PS Claim 18; Page 6; 113pp; English.
XX The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC promegakaryocytes or megakaryocytes, which results in increased platelet
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11
RESULT 37
AAE18011
ID AAE18011 standard; peptide; 14 AA.
XX
AC AAE18011;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human ligand #31 attached to an adenoviral vector.
XX
KW Human; adenoviral coat protein; non-native ligand; cell-surface receptor;
KW therapy; anti-tumour agent; tumour necrosis factor; cancer; brain; lung;
KW ovary; breast; prostate.
XX

PF 22-OCT-2002; 2002WO-US033991.
XX
PR 22-OCT-2001; 2001US-0344614P.
PR 19-SEP-2002; 2002US-0412455P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Rader C, Sinha SC, Lerner R;
XX
DR WPI; 2003-636673/60.
XX
XX Antibody targeting compound useful e.g. for diagnostic immunoassays and
PT treating microbial diseases comprises targeting or biological agent
PT covalently linked to combining site of the antibody.
XX
PS Example 7; Page 62; 56pp; English.
XX
CC The present sequence is that of thrombopoietin (TPO) mimetic peptide
CC AF12505, which mimics the activity of recombinant TPO. The invention
CC provides antibody targeting compounds that are used to reprogram the
CC specificity of an antibody. The antibody targeting compound is linked to
CC the combining site of the antibody, such that the modified antibody takes
CC on the binding specificity of the targeting agent. In an example from the
CC invention, a TPO receptor targeting antibody compound was prepared by
CC covalently linking peptide AB12505 to aldolase monoclonal antibody 38C2.
CC The TPO receptor targeting antibody compound can be used to treat
CC thrombocytopaenia resulting from chemotherapy and bone marrow
CC transplantation
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11

RESULT 40
ADC33697
ID ADC33697 standard; peptide; 14 AA.
XX
AC ADC33697;
XX
DT 18-DEC-2003 (first entry)
XX
DE Erythropoietin receptor/erythropoietin consensus peptide SEQ ID NO:1.
XX
KW chimeric retrovirus envelope protein; ecotropic envelope protein;
KW cytostatic; gene therapy; cancer.
XX
OS Synthetic.
XX
PN WO2003076596-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US007323.
XX
PR 08-MAR-2002; 2002US-0362655P.
XX
PA (UYMA-) UNIV MASSACHUSETTS.
XX
PI Green MR, Gollan TJ;
XX
XX WPI; 2003-722332/68.
DR
XX
PT New chimeric retrovirus envelope protein comprising an ecotropic envelope
PT protein and a heterologous short peptide ligand inserted within the
PT ecotropic envelope protein useful for treating cancer.
XX

PS Disclosure; SEQ ID NO 1; 42pp; English.
XX
CC The present invention describes a chimeric retrovirus envelope protein
CC (I) comprising an ecotropic envelope protein and a heterologous short
CC peptide ligand inserted within the ecotropic envelope protein. Also
CC described: (1) a nucleic acid molecule comprising a sequence encoding the
CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic
CC acid sequence encoding the chimeric envelope protein; (3) a recombinant
CC retroviral particle comprising a chimeric envelope protein comprising a
CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)
CC identifying a nucleic acid sequence encoding the chimeric envelope
CC protein that alters viral tropism; (5) delivering a nucleic acid sequence
CC to a cell; and (6) treating cancer. (I) has cytostatic activity and can
CC be used in gene therapy. The chimeric retrovirus envelope protein is
CC useful for treating cancer, which comprises providing a cancer cell, e.g.
CC human cancer cell and infecting the cancer cell with a virus, e.g.
CC retrovirus comprising the chimeric envelope protein comprising a
CC heterologous short peptide ligand and a therapeutically useful gene, e.g.
CC encoding thymidine kinase. The present sequence represents an
CC erythropoietin receptor/erythropoietin consensus peptide, which is given
CC in the exemplification of the present invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11

RESULT 41
ADJ73004
ID ADJ73004 standard; peptide; 14 AA.
XX
AC ADJ73004;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 458.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 458; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 42
ADN59652
ID ADN59652 standard; peptide; 14 AA.
XX
AC ADN59652;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thrombopoietin mimetic peptide (TMP), seq id 1.
XX
KW Haemostatic; antianaemic; immunosuppressive; platelet;
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;
KW autoimmune haemolytic anaemia; Hughe's syndrome;
KW lupoid thrombocytopaenia.
XX
OS Homo sapiens.
XX
PN WO2003031589-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032552.
XX
PR 11-OCT-2001; 2001US-0328666P.
PR 10-OCT-2002; 2002US-00269806.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Sitney KC, Hartley C;
XX
DR WPI; 2003-403101/38.
XX
PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
PT which stimulate the production of platelets and/or the production of
PT platelet precursors, useful for treating thrombocytopenia.
XX
PS Disclosure; SEQ ID NO 1; 126pp; English.
XX
CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of
CC platelets and/or the production of platelet precursors, is new. Further
CC disclosed is a composition of matter (II) that binds to an mpl receptor,
CC and a pharmaceutical composition comprising (II) and a carrier. The
CC pharmaceutical composition of the invention is useful for treating
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or

CC platelets in a patient. The TMP of the invention is useful for treating
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,
CC autoimmune haemolytic anaemia, Hughe's syndrome and lupoid
CC thrombocytopaenia. The TMP of the invention is also useful for
CC maintaining the viability or storage life of platelets and/or
CC megakaryocytes and its derived cells. The compounds demonstrate an
CC improved ability to bind to and/or trigger transmembrane signal through,
CC i.e. activating, the mpl receptor the compounds have superior
CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in
CC vitro, the production of platelets and/or megakaryocytopoietic activity,
CC i.e. the ability to stimulate, in vivo and in vitro, the production of
CC platelet precursors. Further, certain of the compounds also exhibit
CC superior therapeutic properties, such as improved plasma half-life,
CC biological activity and in vivo circulation time. The current sequence
CC represents a TMP of the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 43
ADJ52639
ID ADJ52639 standard; peptide; 14 AA.
XX
AC ADJ52639;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID458.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 458; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 2 EGPTLRQWLA 11

RESULT 44
ADJ51600
ID ADJ51600 standard; peptide; 14 AA.

AC ADJ51600;

DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID458.

DE CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX Unidentified.
OS Synthetic.

XX WO2004002424-A2.

XX 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 14; SEQ ID NO 458; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | | |
Db 2 EGPTLRQWLA 11

RESULT 45

ADL27293

ID ADL27293 standard; protein; 14 AA.

XX AC ADL27293;

XX 03-JUN-2004 (first entry)

XX Amino acid sequence of a thrombopoietin agonist peptide.

XX fusion protein; C4bp; alpha chain; systemic lupus erythematosus.

XX Homo sapiens.

OS Synthetic.

XX WO2004020639-A2.

XX 11-MAR-2004.

XX 12-AUG-2003; 2003WO-EP008928.

XX 14-AUG-2002; 2002EP-00292043.

XX (AVID-) AVIDIS SA.

XX Garnier L, Hill F, Julien M;

XX WPI; 2004-239202/22.

XX Obtaining a recombinant fusion protein, useful for treating lupus,
PT comprises providing a prokaryotic host cell carrying a nucleic acid
PT encoding the recombinant protein operably linked to a promoter functional
PT in the prokaryotic cell.

XX Claim 8; Page 48; 69pp; English.

XX The specification describes a method for obtaining a recombinant fusion
CC protein comprising a scaffold of a C-terminal core protein of C4bp alpha

CC chain, where the recombinant fusion protein is capable of forming
CC multimers in soluble form in a prokaryotic host cell. The method
CC comprises providing a prokaryotic host cell carrying a nucleic acid
CC encoding the recombinant protein operably linked to a promoter functional
CC in the prokaryotic cell, culturing the host cell under conditions where
CC the recombinant protein is expressed, and recovering the recombinant
CC protein where the protein is recovered in multimeric form without
CC performing a scaffold refolding step. The protein is useful for treating
CC systemic lupus erythematosus. The present sequence represents a
CC thrombopoietin agonist peptide, which is used to produce fusion proteins
CC of the invention.

XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 46
ADM72503
ID ADM72503 standard; peptide; 14 AA.
XX
AC ADM72503;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 14 /note= "sarcosine"
FT
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
PT Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.

CC The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.

CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the incidence of secondary failure
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 47
ADM72483
ID ADM72483 standard; peptide; 14 AA.
XX
AC ADM72483;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
PT Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.

XX The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the time of platelet and/or
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment, by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11

RESULT 48
ADM72495
ID ADM72495 standard; peptide; 14 AA.
XX
AC ADM72495;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 13
FT /note= "p-amino-Phe"
FT
XX
PN WO2004026332-A1.
XX
XX 01-APR-2004.
PD
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.

PR 18-SEP-2002; 2002US-0411779P.
XX (THRE-) 3-DIMENSIONAL PHARM INC.
PA
XX Kaushansky K, Macdonald BR;
PI
XX WPI; 2004-283153/26.
DR
XX Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
PS Disclosure; Fig 2; 32pp; English.
XX
CC The invention relates to a method (M1) for increasing haematopoietic stem
CC cell production in a subject which involves administering a
CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the time of platelet and/or
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db |||||||
2 EGPTLRQWLA 11

RESULT 49
ADM72487
ID ADM72487 standard; peptide; 14 AA.
XX
AC ADM72487;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

OS Synthetic.

XX Key Location/Qualifiers

XX Modified-site 14

FT /label= bAla

FT /note= "beta-alanine"

XX

PN WO2004026332-A1.

XX

PD 01-APR-2004.

XX

PF 18-SEP-2003; 2003WO-US029701.

XX

PR 18-SEP-2002; 2002US-0411700P.

PR 18-SEP-2002; 2002US-0411779P.

XX

PA (THRE-) 3-DIMENSIONAL PHARM INC.

XX

PI Kaushansky K, Macdonald BR;

XX

DR WPI; 2004-283153/26.

XX

PT Increasing hematopoietic stem cell production in subject, useful in

PT reducing the incidence of delayed primary engraftment, comprises

PT administering a Thrombopoietin mimetic compound e.g., a peptide to a

PT subject.

XX

PS Disclosure; Fig 2; 32pp; English.

XX

CC The invention relates to a method (M1) for increasing haematopoietic stem

CC cell production in a subject which involves administering a

CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is

CC another method (M2) of providing haematopoietic stem cells to a subject

CC which involves administering a TPO mimetic compound to a subject to

CC enhance expansion of a stem cell population within bone marrow and/or

CC mobilize stem cells in peripheral circulation, harvesting one or more of

CC the bone marrow stem cells or the stem cells in the peripheral

CC circulation, and transplanting the harvested stem cells into the subject.

CC A method (M3) is also provided for reducing a time to engraftment

CC following reinfusion of stem cells in a subject, involves administering a

CC TPO mimetic compound to the subject, enhancing the expansion of the stem

CC cell population within bone marrow and/or mobilizing the stem cells in

CC peripheral circulation, harvesting one or more of the bone marrow stem

CC cells or one or more of the stem cells in the peripheral circulation, and

CC transplanting the one or more harvested stem cells into the subject. TPO

CC mimetic compounds are disclosed as peptides, including cyclic or modified

CC peptides. (M1) is useful for increasing haematopoietic stem cell

CC production in a subject e.g., human. (M3) is useful for reducing time to

CC engraftment following reinfusion of stem cells, reducing the incidence of

CC delayed primary engraftment, reducing the incidence of secondary failure

CC of platelet production and reducing the time of platelet and/or

CC neutrophil engraftment following reinfusion of stem cells in a subject.

CC (M1) is also useful for increasing the number of stem cells from a donor

CC whose cells are then used for rescue of recipient subject. Also useful in

CC the treatment of thrombocytopenia. (M1) enables transplantation to

CC proceed in patients who would not otherwise be considered as candidates

CC because of unacceptably high risk of failed engraftment, reduces the

CC number of aphereses required to generate a minimum acceptable harvest,

CC reduces the incidence of primary and secondary failure of engraftment by

CC increasing the number of haematopoietic stem cells (HSCs) available for

CC transplantation and reduces the time required for primary engraftment.

CC The present sequence represents an example of TPO mimetic peptide

CC fragment.

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;

Best Local Similarity 100.0%; Pred. NO. 0.006;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 EGPTLRQWLA 11

RESULT 50

ADM72497

ID ADM72497 standard; peptide; 14 AA.

XX

AC ADM72497;

XX

DT 17-JUN-2004 (first entry)

XX

DE TPO mimetic peptide fragment.

XX

XW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;

KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 13

FT /note= "Ac-Lys"

XX

PN WO2004026332-A1.

XX

PD 01-APR-2004.

XX

PF 18-SEP-2003; 2003WO-US029701.

XX

PR 18-SEP-2002; 2002US-0411700P.

PR 18-SEP-2002; 2002US-0411779P.

XX

PA (THRE-) 3-DIMENSIONAL PHARM INC.

XX

PI Kaushansky K, Macdonald BR;

XX

DR WPI; 2004-283153/26.

XX

PT Increasing hematopoietic stem cell production in subject, useful in

PT reducing the incidence of delayed primary engraftment, comprises

PT administering a Thrombopoietin mimetic compound e.g., a peptide to a

PT subject.

XX

PS Disclosure; Fig 2; 32pp; English.

XX

CC The invention relates to a method (M1) for increasing haematopoietic stem

CC cell production in a subject which involves administering a

CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is

CC another method (M2) of providing haematopoietic stem cells to a subject

CC which involves administering a TPO mimetic compound to a subject to

CC enhance expansion of a stem cell population within bone marrow and/or

CC mobilize stem cells in peripheral circulation, harvesting one or more of

CC the bone marrow stem cells or the stem cells in the peripheral

CC circulation, and transplanting the harvested stem cells into the subject.

CC A method (M3) is also provided for reducing a time to engraftment

CC following reinfusion of stem cells in a subject, involves administering a

CC TPO mimetic compound to the subject, enhancing the expansion of the stem

CC cell population within bone marrow and/or mobilizing the stem cells in

CC peripheral circulation, harvesting one or more of the bone marrow stem

CC cells or one or more of the stem cells in the peripheral circulation, and

CC transplanting the one or more harvested stem cells into the subject. TPO

CC mimetic compounds are disclosed as peptides, including cyclic or modified

CC peptides. (M1) is useful for increasing haematopoietic stem cell

CC production in a subject e.g., human. (M3) is useful for reducing time to

CC engraftment following reinfusion of stem cells, reducing the incidence of

CC delayed primary engraftment, reducing the incidence of secondary failure

CC of platelet production and reducing the time of platelet and/or

CC neutrophil engraftment following reinfusion of stem cells in a subject.

CC (M1) is also useful for increasing the number of stem cells from a donor

CC whose cells are then used for rescue of recipient subject. Also useful in

CC the treatment of thrombocytopenia. (M1) enables transplantation to

CC proceed in patients who would not otherwise be considered as candidates

CC because of unacceptably high risk of failed engraftment, reduces the

CC number of aphereses required to generate a minimum acceptable harvest,

CC reduces the incidence of primary and secondary failure of engraftment by

CC increasing the number of haematopoietic stem cells (HSCs) available for

CC transplantation and reduces the time required for primary engraftment.

CC The present sequence represents an example of TPO mimetic peptide

CC fragment.

XX

Qy 1 EGPTLRQWLA 10

|||||||

CC reduces the incidence of primary and secondary failure of engraftment by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | | | | |
Db 2 EGPTLRQWLA 11

Search completed: May 12, 2006, 10:36:47
Job time : 77.7496 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-62
Perfect score: 56
Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

```
Database :      PIR_80:*
1:  _pir1:
2:  _pir2:
3:  _pir3:
4:  _pir4:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	42	75.0	1019	2	T11560	pol polyprotein -
2	40	71.4	296	2	AG0147	probable membrane
3	39	69.6	131	2	S74539	hypothetical prote
4	39	69.6	217	2	S46354	pol polyprotein -
5	39	69.6	473	2	E84853	hypothetical prote
6	39	69.6	656	2	S30484	pol polyprotein -
7	39	69.6	656	2	S30483	pol polyprotein -
8	39	69.6	1034	1	GNLJCA	HIV-1 retropepsin
9	39	69.6	1035	1	GNLJGG	HIV-1 retropepsin
10	39	69.6	1036	1	GNLJG2	HIV-1 retropepsin
11	39	69.6	1055	1	GNLJST	HIV-1 retropepsin
12	39	69.6	1055	2	S53092	pol polyprotein -
13	38	67.9	134	2	B75468	hypothetical prote
14	38	67.9	1712	1	CGHU2B	collagen alpha 2(I
15	37.5	67.0	333	2	A36925	transcription acti
16	37	66.1	306	2	D70601	UTP-glucose-1-phos
17	37	66.1	306	2	T45453	UTP-glucose-1-phos
18	37	66.1	547	2	T36550	hypothetical prote
19	37	66.1	777	2	A87309	hypothetical prote
20	37	66.1	816	2	A71006	hypothetical prote
21	37	66.1	825	2	JC4163	DNA-binding protei
22	37	66.1	1039	2	S46347	pol polyprotein -
23	37	66.1	1058	2	S08436	regulator protein
24	37	66.1	1083	2	S54293	HIV-1 retropepsin
25	36	64.3	151	2	S63748	hypothetical prote
26	36	64.3	200	2	T23485	probable glutathio
27	36	64.3	207	2	T37464	probable transposa
28	36	64.3	278	2	T45494	hypothetical prote
29	36	64.3	312	2	F86876	hypothetical prote

103 34 60.7 458 2 B83372 hypothetical prote
104 34 60.7 463 2 S27491 hypothetical prote
105 34 60.7 469 2 AD1926 hypothetical prote
106 34 60.7 470 2 AD0888 Sufi protein [impo
107 34 60.7 470 2 E91116 suppressor of ftsI
108 34 60.7 470 2 E85961 suppressor of ftsI
109 34 60.7 470 2 G65088 sufi protein precu
110 34 60.7 571 2 AI0506 probable sulfatase
111 34 60.7 594 2 G83878 L-lactate permease
112 34 60.7 750 2 A97501 topoisomerase iv c
113 34 60.7 750 2 AE2719 topoisomerase IV s
114 34 60.7 868 2 C81200 aconitate hydratase
115 34 60.7 868 2 H81775 aconitate hydratase
116 34 60.7 955 2 T48515 hypothetical prote
117 34 60.7 1009 2 S28081 pol polyprotein -
118 34 60.7 1054 1 GNLJG5 HIV-1 retropepsin
119 34 60.7 1061 1 GNLJG4 HIV-1 retropepsin
120 34 60.7 1123 2 T28139 PK4 protein kinase
121 34 60.7 1155 2 AC2426 adenylate cyclase
122 34 60.7 1168 2 T30935 reverse transcript
123 34 60.7 1707 2 A33526 collagen alpha 2(I
124 34 60.7 2329 2 S44625 C50C3.6 protein -
125 34 60.7 2363 2 T38841 probable pre-mRNA
126 34 60.7 3433 1 GNVVKV genome polyprotein
127 33.5 59.8 209 2 B87627 hypothetical prote
128 33 58.9 39 2 I37555 homeobox - human (
129 33 58.9 75 2 S03065 ser protein - huma
130 33 58.9 85 2 AH1010 conserved hypothet
131 33 58.9 129 2 S63743 HIV-1 retropepsin
132 33 58.9 130 2 S63754 HIV-1 retropepsin
133 33 58.9 132 2 S63733 HIV-1 retropepsin
134 33 58.9 136 2 S63746 HIV-1 retropepsin
135 33 58.9 140 2 S63737 HIV-1 retropepsin
136 33 58.9 141 2 S63750 HIV-1 retropepsin
137 33 58.9 142 2 AF0961 heat shock protein
138 33 58.9 143 2 S63735 HIV-1 retropepsin
139 33 58.9 145 2 S63745 HIV-1 retropepsin
140 33 58.9 149 2 S63734 HIV-1 retropepsin
141 33 58.9 151 2 S63731 HIV-1 retropepsin
142 33 58.9 153 2 A96751 hypothetical prote
143 33 58.9 154 2 AC0496 heat shock protein
144 33 58.9 160 2 S63753 HIV-1 retropepsin
145 33 58.9 162 2 S63732 HIV-1 retropepsin
146 33 58.9 170 2 E83871 hypothetical prote
147 33 58.9 176 2 S63747 HIV-1 retropepsin
148 33 58.9 190 2 A82343 conserved hypothet
149 33 58.9 208 2 AD2808 uracil-DNA glycosy
150 33 58.9 213 2 AC3389 hypothetical prote
151 33 58.9 213 2 AH2253 vif protein - simi
152 33 58.9 214 1 ASLJG5 vif protein - simi
153 33 58.9 217 1 B97587 hypothetical prote
154 33 58.9 217 2 B97587 RNA-directed DNA p
155 33 58.9 219 2 S32096 RNA-directed DNA p
156 33 58.9 219 2 S32157 RNA-directed DNA p
157 33 58.9 219 2 S32080 RNA-directed DNA p
158 33 58.9 219 2 S32062 RNA-directed DNA p
159 33 58.9 219 2 S32047 RNA-directed DNA p
160 33 58.9 219 2 S32065 RNA-directed DNA p
161 33 58.9 219 2 S32122 RNA-directed DNA p
162 33 58.9 219 2 S32060 RNA-directed DNA p
163 33 58.9 219 2 S32118 RNA-directed DNA p
164 33 58.9 219 2 S32057 RNA-directed DNA p
165 33 58.9 219 2 S32048 RNA-directed DNA p
166 33 58.9 219 2 S32094 RNA-directed DNA p
167 33 58.9 219 2 S32061 RNA-directed DNA p
168 33 58.9 219 2 S32084 RNA-directed DNA p
169 33 58.9 219 2 S32152 RNA-directed DNA p
170 33 58.9 219 2 S32089 RNA-directed DNA p
171 33 58.9 219 2 S32073 RNA-directed DNA p
172 33 58.9 219 2 S32049 RNA-directed DNA p
173 33 58.9 219 2 S32120 RNA-directed DNA p
174 33 58.9 219 2 S32086 RNA-directed DNA p
175 33 58.9 219 2 S32133 RNA-directed DNA p

176 33 58.9 219 2 S32081 RNA-directed DNA p
177 33 58.9 219 2 S32139 RNA-directed DNA p
178 33 58.9 219 2 S32078 RNA-directed DNA p
179 33 58.9 219 2 S32066 RNA-directed DNA p
180 33 58.9 219 2 S32126 RNA-directed DNA p
181 33 58.9 219 2 S32074 RNA-directed DNA p
182 33 58.9 219 2 S32053 RNA-directed DNA p
183 33 58.9 219 2 S32088 RNA-directed DNA p
184 33 58.9 219 2 S32072 RNA-directed DNA p
185 33 58.9 219 2 S32070 RNA-directed DNA p
186 33 58.9 219 2 S32087 RNA-directed DNA p
187 33 58.9 219 2 S32076 RNA-directed DNA p
188 33 58.9 219 2 S32092 RNA-directed DNA p
189 33 58.9 219 2 S32160 RNA-directed DNA p
190 33 58.9 219 2 S32054 RNA-directed DNA p
191 33 58.9 219 2 S32135 RNA-directed DNA p
192 33 58.9 219 2 S32159 RNA-directed DNA p
193 33 58.9 219 2 S32069 RNA-directed DNA p
194 33 58.9 219 2 S32063 RNA-directed DNA p
195 33 58.9 219 2 S32128 RNA-directed DNA p
196 33 58.9 219 2 S32085 RNA-directed DNA p
197 33 58.9 219 2 S32064 RNA-directed DNA p
198 33 58.9 219 2 S32056 RNA-directed DNA p
199 33 58.9 219 2 S32059 RNA-directed DNA p
200 33 58.9 219 2 S32131 RNA-directed DNA p
201 33 58.9 219 2 S32095 RNA-directed DNA p
202 33 58.9 219 2 S32071 RNA-directed DNA p
203 33 58.9 219 2 S32127 RNA-directed DNA p
204 33 58.9 219 2 S32134 RNA-directed DNA p
205 33 58.9 219 2 S32093 RNA-directed DNA p
206 33 58.9 219 2 S32075 RNA-directed DNA p
207 33 58.9 219 2 S32079 RNA-directed DNA p
208 33 58.9 219 2 S32119 RNA-directed DNA p
209 33 58.9 219 2 S32138 RNA-directed DNA p
210 33 58.9 219 2 S32129 RNA-directed DNA p
211 33 58.9 219 2 S32097 RNA-directed DNA p
212 33 58.9 219 2 S32117 RNA-directed DNA p
213 33 58.9 219 2 S32077 RNA-directed DNA p
214 33 58.9 219 2 S32098 RNA-directed DNA p
215 33 58.9 219 2 S32051 RNA-directed DNA p
216 33 58.9 219 2 A46097 GPI-anchor biosynt
217 33 58.9 219 2 S32132 RNA-directed DNA p
218 33 58.9 219 2 S32058 RNA-directed DNA p
219 33 58.9 224 2 S32140 RNA-directed DNA p
220 33 58.9 224 2 F70614 probable malonyl c
221 33 58.9 235 2 C83822 hypothetical prote
222 33 58.9 260 2 E81194 biotin synthesis p
223 33 58.9 262 2 I40221 divIB protein - Ba
224 33 58.9 264 2 AG2095 hypothetical prote
225 33 58.9 302 2 S75227 regulatory protein
226 33 58.9 311 1 RGECK transcription acti
227 33 58.9 311 2 AH0867 positive regulator
228 33 58.9 311 2 C85936 positive regulator
229 33 58.9 311 2 H91090 probable dehydroge
230 33 58.9 313 2 T35826 probable transposa
231 33 58.9 327 2 E82277 glyceraldehyde-3-p
232 33 58.9 331 2 B48445 hypothetical prote
233 33 58.9 336 2 E72389 homeobox protein c
234 33 58.9 337 2 T27635 transcription regu
235 33 58.9 338 2 AH3321 probable exported
236 33 58.9 341 2 AG0195 conserved hypothet
237 33 58.9 352 2 G83636 hypothetical prote
238 33 58.9 364 2 C84187 ubiquinol-cytochro
239 33 58.9 380 2 A71390 hypothetical prote
240 33 58.9 383 2 T20572 corrinoide/iron-sul
241 33 58.9 389 2 B69096 hypothetical prote
242 33 58.9 410 2 H86290 N-ethylmellamine ch
243 33 58.9 416 2 E69374 histidyl-tRNA synt
244 33 58.9 431 2 A81150 conserved hypothet
245 33 58.9 451 2 AF3053 hypothetical prote
246 33 58.9 451 2 G98232 multidrug resistan
247 33 58.9 472 2 AI3367 probable MFS trans
248 33 58.9 477 2 H83588

249	33	58.9	499	2	H83254	probable MFS trans	322	32	57.1	218	2	JC7220	nuclear protein SR
250	33	58.9	510	2	S41943	cellulose 1,4-beta	323	32	57.1	220	2	D82139	conserved hypothet
251	33	58.9	511	2	S44716	cellulose 1,4-beta	324	32	57.1	221	2	B36881	MPM2-reactive phos
252	33	58.9	536	1	SYCEB	2,3-dihydroxybenzo	325	32	57.1	222	2	B90888	hypothetical prote
253	33	58.9	536	2	E85558	2,3-dihydroxybenzo	326	32	57.1	229	2	JC7219	nuclear protein SR
254	33	58.9	536	2	A99708	2,3-dihydroxybenzo	327	32	57.1	244	2	A82316	hypothetical prote
255	33	58.9	540	2	S41942	cellulose 1,4-beta	328	32	57.1	249	2	E87575	ABC transporter, A
256	33	58.9	542	1	A54963	transcription fact	329	32	57.1	252	2	C84522	22 kDa peroxisomal
257	33	58.9	544	2	A72459	probable glutamyl-	330	32	57.1	254	2	T35994	probable ABC-type
258	33	58.9	559	2	B47175	reverse transcript	331	32	57.1	255	2	A45881	MHC class II histo
259	33	58.9	559	2	A47175	reverse transcript	332	32	57.1	269	2	S73999	hypothetical prote
260	33	58.9	567	2	T16105	hypothetical prote	333	32	57.1	274	2	T05509	hypothetical prote
261	33	58.9	582	2	T16104	hypothetical prote	334	32	57.1	278	2	F84127	hypothetical prote
262	33	58.9	586	2	B84271	glutamyl-tRNA synt	335	32	57.1	285	2	G71337	probable dimethyla
263	33	58.9	589	2	F87626	sensor histidine k	336	32	57.1	288	2	S73016	H+-transporting tw
264	33	58.9	600	2	C83221	transport protein	337	32	57.1	291	2	T10966	6-phosphogluconate
265	33	58.9	618	2	AI0171	probable exported	338	32	57.1	293	2	T35157	6-phosphogluconate
266	33	58.9	625	2	T40742	hypothetical integ	339	32	57.1	302	2	T36030	probable fructokin
267	33	58.9	629	2	T16767	hypothetical prote	340	32	57.1	309	2	F83434	translocation prot
268	33	58.9	633	2	B70946	NADH2 dehydrogenas	341	32	57.1	310	2	JC7853	L-fucose-specific
269	33	58.9	709	2	S75212	comE protein - Syn	342	32	57.1	312	2	S76507	hypothetical prote
270	33	58.9	719	2	B95325	conserved hypothet	343	32	57.1	326	2	C75350	probable UV damage
271	33	58.9	739	2	T45429	polyphosphate kina	344	32	57.1	327	2	H82736	hypothetical prote
272	33	58.9	742	2	E70673	probable ppk prote	345	32	57.1	331	2	A72514	hypothetical prote
273	33	58.9	772	2	E82592	transcription-rela	346	32	57.1	355	2	T20825	hypothetical prote
274	33	58.9	838	2	B83150	probable Arp-depen	347	32	57.1	362	2	A64212	protein serine/thr
275	33	58.9	902	2	T01668	pol polyprotein -	348	32	57.1	368	2	G65119	hypothetical 40.4
276	33	58.9	912	2	S33980	pol polyprotein -	349	32	57.1	373	2	B95291	conserved hypothet
277	33	58.9	924	2	E71476	alanine-tRNA ligas	350	32	57.1	393	2	B85992	probable transport
278	33	58.9	994	2	AC3480	DNA-directed DNA p	351	32	57.1	394	2	C82439	peptide methionine
279	33	58.9	1000	2	T30280	hypothetical prote	352	32	57.1	396	1	XNECD	aspartate transami
280	33	58.9	1002	1	GNLJND	HIV-1 retropepsin	353	32	57.1	396	2	AD0616	aspartate aminotra
281	33	58.9	1002	2	S54378	pol polyprotein -	354	32	57.1	396	2	A85619	aspartate aminotra
282	33	58.9	1003	1	GNVWL	HIV-1 retropepsin	355	32	57.1	396	2	C90755	aspartate aminotra
283	33	58.9	1003	1	B44001	HIV-1 retropepsin	356	32	57.1	399	2	F91146	probable transport
284	33	58.9	1003	1	GNVWA2	HIV-1 retropepsin	357	32	57.1	400	2	T18570	probable malate de
285	33	58.9	1003	2	T09440	HIV-1 retropepsin	358	32	57.1	408	2	T50876	hypothetical membr
286	33	58.9	1012	1	GNVWVL	pol polyprotein -	359	32	57.1	409	2	T47118	thiamine pyridinyl
287	33	58.9	1015	1	GNVWH3	HIV-1 retropepsin	360	32	57.1	418	2	S46315	aspartate transami
288	33	58.9	1027	1	GNLJSI	HIV-1 retropepsin	361	32	57.1	424	2	A35861	interferon consens
289	33	58.9	1073	2	I51055	HIV-1 retropepsin	362	32	57.1	425	2	A45064	interferon consens
290	33	58.9	1094	2	F70697	recombination acti	363	32	57.1	428	2	B71403	hypothetical prote
291	33	58.9	1124	1	GNLJFP	probable arabinosy	364	32	57.1	438	2	G87337	membrane protein,
292	33	58.9	1124	2	B45557	HIV-1 retropepsin	365	32	57.1	454	2	C82682	glutamate-cysteine
293	33	58.9	1124	2	S23820	pol polyprotein -	366	32	57.1	456	2	T06136	aspartate transami
294	33	58.9	1233	1	P3XRD3	major core protein	367	32	57.1	473	2	AI0587	deoxyribodipyrimid
295	33	58.9	1733	2	D70887	probable polyketid	368	32	57.1	473	2	S22321	deoxyribodipyrimid
296	33	58.9	1784	2	E86921	polyketide synthas	369	32	57.1	475	2	T01352	hypothetical prote
297	33	58.9	2399	2	H71879	toxin-like outer m	370	32	57.1	482	2	D75346	glutamyl-tRNA(Gln)
298	33	58.9	2529	2	B64635	toxin-like outer m	371	32	57.1	489	2	AI1276	multidrug-efflux t
299	33	58.9	3430	1	GNVWV	genome polyprotein	372	32	57.1	489	2	AI1539	multidrug-efflux t
300	32.5	58.0	235	2	S55883	CCHH finger protei	373	32	57.1	493	2	ADI398	drug-export protei
301	32.5	58.0	245	2	A84279	hypothetical prote	374	32	57.1	493	2	AG1773	drug-export protei
302	32.5	58.0	480	2	T36822	probable transcrip	375	32	57.1	504	2	A49467	occludin - chicken
303	32	57.1	84	2	AB2301	hypothetical prote	376	32	57.1	518	2	AD2315	hypothetical prote
304	32	57.1	111	2	H95908	conserved hypothet	377	32	57.1	519	2	D82536	conserved hypothet
305	32	57.1	117	2	AB2757	hypothetical prote	378	32	57.1	525	2	D90048	glutamate synthase
306	32	57.1	123	2	T49336	hypothetical prote	379	32	57.1	525	2	C69794	glutamate synthase
307	32	57.1	125	2	F81197	holo-(acyl-carrier	380	32	57.1	532	1	E69821	multidrug resistan
308	32	57.1	147	2	T38472	hypothetical prote	381	32	57.1	540	2	B45665	adult-specific 61.
309	32	57.1	183	2	AC0767	dTDP-4-dehydrodham	382	32	57.1	559	2	AI0135	DNA repair protein
310	32	57.1	183	2	S15302	hypothetical prote	383	32	57.1	560	2	S65148	tRNA adenyllyltrans
311	32	57.1	193	2	T35847	probable carbonic	384	32	57.1	568	2	T05218	hypothetical prote
312	32	57.1	195	2	F91171	probable phosphopa	385	32	57.1	573	2	I60247	SEC23 protein homo
313	32	57.1	195	2	F86017	hypothetical prote	386	32	57.1	575	2	AI1996	hypothetical prote
314	32	57.1	195	2	S47694	hypothetical 21.8K	387	32	57.1	594	2	B82994	probable secretion
315	32	57.1	196	2	AI0502	molybdopterin bios	388	32	57.1	601	2	T36323	probable membrane
316	32	57.1	198	2	B84382	hypothetical prote	389	32	57.1	606	2	G64659	flagellar hook-ass
317	32	57.1	207	2	B75327	hypothetical prote	390	32	57.1	610	2	AE2436	penicillin-binding
318	32	57.1	209	2	A83849	alpha-ribazole-5'-	391	32	57.1	615	2	AI0615	probable exported
319	32	57.1	210	2	G85729	hypothetical prote	392	32	57.1	615	2	H90754	probable amidase l
320	32	57.1	214	2	S07989	vif protein - simi	393	32	57.1	615	2	F85618	probable amidase y
321	32	57.1	214	2	T11561	vif protein - simi	394	32	57.1	615	2	D64832	ycbB protein precu

395 32 57.1 623 2 T48859 disease resistance
396 32 57.1 623 2 T06674 hypothetical prote
397 32 57.1 629 2 T06675 hypothetical prote
398 32 57.1 664 2 G89894 protein kinase [im
399 32 57.1 691 2 T44543 probable bacteriop
400 32 57.1 715 2 S38051 DOA1 protein - yea
401 32 57.1 725 2 A11544 conserved hypothet
402 32 57.1 728 2 D86278 hypothetical prote
403 32 57.1 755 2 D75598 photoreceptor - De
404 32 57.1 761 2 T09052 hypothetical prote
405 32 57.1 763 2 E96571 hypothetical prote
406 32 57.1 765 2 T09574 transport protein
407 32 57.1 775 2 B64319 carbon-monoxide de
408 32 57.1 798 2 C98069 primosomal replica
409 32 57.1 808 2 E64914 dimethylsulfoxide
410 32 57.1 808 2 E90915 probable oxidoredu
411 32 57.1 808 2 B85764 probable oxidoredu
412 32 57.1 916 2 H72372 excinuclease ABC c
413 32 57.1 916 2 G75417 SNF2/Rad54 helicase
414 32 57.1 921 2 A40597 DNA-directed DNA p
415 32 57.1 942 2 S75598 poly(A) polymerase
416 32 57.1 943 2 T34847 probable transcrip
417 32 57.1 1003 2 T13856 ker protein - frui
418 32 57.1 1016 1 S40838 formate dehydrogen
419 32 57.1 1016 2 D91231 formate dehydrogen
420 32 57.1 1019 2 T00117 dve protein - frui
421 32 57.1 1099 1 S31926 myosin IB heavy ch
422 32 57.1 1099 2 A59300 myosin-If - mouse
423 32 57.1 1102 2 S55100 hypothetical prote
424 32 57.1 1170 2 S03308 cell surface glyco
425 32 57.1 1175 2 T46124 hypothetical prote
426 32 57.1 1268 2 S52781 neurocan - mouse
427 32 57.1 1333 2 S30356 CDC25 protein homo
428 32 57.1 1369 2 T03104 tegument protein h
429 32 57.1 1541 2 S46686 hypothetical prote
430 32 57.1 1545 2 T26589 hypothetical prote
431 32 57.1 1616 2 T17884 S-layer protein -
432 32 57.1 1620 2 E83261 conserved hypothet
433 32 57.1 2108 2 H70819 probable polyketid
434 31.5 56.2 332 2 F83600 conserved hypothet
435 31.5 56.2 909 2 D87434 phosphoenolpyruvat
436 31.5 56.2 911 2 AD0959 Two-component sens
437 31.5 56.2 1080 2 T03964 probable ubiquitin
438 31 55.4 77 2 A83468 hypothetical prote
439 31 55.4 85 2 B91147 hypothetical prote
440 31 55.4 85 2 F85992 hypothetical prote
441 31 55.4 85 2 C65120 hypothetical 10.0
442 31 55.4 93 2 E70967 hypothetical prote
443 31 55.4 108 2 G84522 similar to gibbere
444 31 55.4 108 2 T49731 hypothetical prote
445 31 55.4 110 2 G90584 50S ribosomal prot
446 31 55.4 111 1 R5YM22 ribosomal protein
447 31 55.4 111 2 G70521 hypothetical prote
448 31 55.4 112 2 S75152 hypothetical prote
449 31 55.4 113 2 S10612 ribosomal protein
450 31 55.4 113 2 H69696 ribosomal protein
451 31 55.4 114 2 A10294 probable membrane
452 31 55.4 117 2 H90021 50S ribosomal prot
453 31 55.4 118 2 AC1403 ribosomal protein
454 31 55.4 118 2 AB1779 ribosomal protein
455 31 55.4 125 2 E82378 hypothetical prote
456 31 55.4 144 2 B91207 heat shock protein
457 31 55.4 144 2 G65170 heat shock protein
458 31 55.4 144 2 D86053 heat shock protein
459 31 55.4 147 2 A70928 hypothetical prote
460 31 55.4 150 2 S37014 transposase (clone
461 31 55.4 153 2 A97524 hypothetical prote
462 31 55.4 153 2 C69193 ribosomal protein
463 31 55.4 155 2 D75147 lsu ribosomal prot
464 31 55.4 155 2 B71187 probable ribosomal
465 31 55.4 162 2 A34429 dihydrofolate redu
466 31 55.4 165 2 G69224 conserved hypothet
467 31 55.4 174 2 F82786 conserved hypothet

31 55.4 178 2 AB2743 hypothetical prote
31 55.4 189 2 F95017 hypothetical prote
31 55.4 189 2 G97890 hypothetical prote
31 55.4 203 2 H95842 probable glutathio
31 55.4 209 2 D95274 hypothetical prote
31 55.4 216 2 D75567 GRP cyclohydrolase
31 55.4 218 2 B71183 hypothetical prote
31 55.4 219 2 E75143 phosphoglycolate p
31 55.4 219 2 B97461 structure of gluta
31 55.4 219 2 AD2679 Glutathione-S-tran
31 55.4 219 2 T38159 probable RNA-bindin
31 55.4 233 2 A83862 initiation of chro
31 55.4 247 2 S18604 triose-phosphate i
31 55.4 249 2 A70840 hypothetical prote
31 55.4 260 2 F81830 conserved hypothet
31 55.4 260 2 C64129 biotin synthesis p
31 55.4 261 2 G84010 hypothetical prote
31 55.4 262 2 B83126 probable transcrip
31 55.4 263 2 A30227 hypothetical prote
31 55.4 264 2 AH0759 propanediol diffus
31 55.4 264 2 AH2202 hypothetical prote
31 55.4 264 2 G83343 hypothetical prote
31 55.4 268 2 H69838 chloride peroxidase
31 55.4 268 2 B42424 chitinase (EC 3.2.
31 55.4 269 2 C84314 hypothetical prote
31 55.4 272 2 T35231 hypothetical prote
31 55.4 281 2 H86159 hypothetical prote
31 55.4 286 2 G70537 hypothetical prote
31 55.4 289 2 JQ0059 glutamyl-tRNA synt
31 55.4 295 2 E75366 hypothetical prote
31 55.4 307 2 B83167 hypothetical prote
31 55.4 326 2 C24430 glyceraldehyde-3-p
31 55.4 329 2 S55489 probable galactosyl
31 55.4 329 2 S73710 adhesin P1 homolog
31 55.4 332 2 T04754 hypothetical prote
31 55.4 335 2 H70545 probable fabH prot
31 55.4 336 1 DEPZG glyceraldehyde-3-p
31 55.4 337 2 A35080 glyceraldehyde-3-p
31 55.4 338 1 DEIS3C glyceraldehyde-3-p
31 55.4 338 2 JQ1287 glyceraldehyde-3-p
31 55.4 338 2 E90998 hypothetical prote
31 55.4 339 2 AG0799 probable transcrip
31 55.4 342 1 S64042 porphobilinogen sy
31 55.4 346 2 D85818 unknown protein en
31 55.4 355 2 AG1877 hypothetical prote
31 55.4 359 2 D83103 probable phospholi
31 55.4 382 2 T11138 ubiquinol-cytochro
31 55.4 382 2 S77130 hypothetical prote
31 55.4 387 2 A10511 probable metabolit
31 55.4 392 2 T51151 probable nuclear D
31 55.4 396 2 S63374 diphosphomevalonat
31 55.4 400 2 C87021 serine-threonine p
31 55.4 406 2 AB1060 arginine deiminase
31 55.4 406 2 A86298 hypothetical prote
31 55.4 407 2 T36404 probable monooxyge
31 55.4 409 2 T37507 aspartate transami
31 55.4 412 2 S76130 hypothetical prote
31 55.4 416 1 JC4952 transcription init
31 55.4 417 2 F97789 ampG protein [impo
31 55.4 422 2 F96826 hypothetical prote
31 55.4 424 2 T25774 hypothetical prote
31 55.4 425 2 A87517 aminotransferase,
31 55.4 427 2 S30558 ribose-phosphate d
31 55.4 427 2 B83001 probable C4-dicarb
31 55.4 430 2 F87472 FMN oxidoreductase
31 55.4 433 2 S51837 glyceraldehyde-3-p
31 55.4 433 2 S51836 glyceraldehyde-3-p
31 55.4 437 1 B29336 ubiquinol-cytochro
31 55.4 449 2 T48511 aspartate transami
31 55.4 450 2 E98303 hypothetical 49.3K

541 31 55.4 450 2 AH2979 nitrilotriacetate
542 31 55.4 454 2 JC7231 thermophilic desul
543 31 55.4 454 2 F87149 probale secreted p
544 31 55.4 475 2 F84261 probable transport
545 31 55.4 477 2 B83476 probable transport
546 31 55.4 483 2 D71439 probable Beta-Amyl
547 31 55.4 491 2 G89860 hypothetical prote
548 31 55.4 496 2 G87546 acid-CoA ligase, p
549 31 55.4 498 2 H85190 probable beta-amyl
550 31 55.4 499 2 H87460 hypothetical prote
551 31 55.4 499 2 S39113 alpha-N-arabinofur
552 31 55.4 503 2 B83490 probable MFS trans
553 31 55.4 511 1 E70391 major facilitator
554 31 55.4 524 2 S75852 hypothetical prote
555 31 55.4 526 2 C83488 probable semialdeh
556 31 55.4 529 2 H75078 ATP-dependent RNA
557 31 55.4 541 2 F84187 phytoene dehydroge
558 31 55.4 548 2 T52556 beta-amylase (EC 3
559 31 55.4 557 2 T44843 glucose-6-phosphat
560 31 55.4 565 1 C72550 methylmalonyl-CoA
561 31 55.4 572 2 T47219 amino acid transpo
562 31 55.4 575 2 C86398 protein T7N9.26 [i
563 31 55.4 589 2 T34878 probable integral
564 31 55.4 610 1 S30234 transcription fact
565 31 55.4 612 2 AI3493 oligopeptide-bindi
566 31 55.4 628 2 S35547 double-strand-brea
567 31 55.4 645 2 G88130 protein F10G7.4 [i
568 31 55.4 651 2 T25953 hypothetical prote
569 31 55.4 664 2 H83962 serine/threonine p
570 31 55.4 750 2 D86245 hypothetical prote
571 31 55.4 762 2 H83415 cis/trans isomeras
572 31 55.4 766 2 E70916 probable bisc prot
573 31 55.4 775 2 AB3402 topoisomerase IV c
574 31 55.4 777 1 G69773 conserved hypothet
575 31 55.4 789 2 S28259 androgen-regulated
576 31 55.4 792 2 A84308 chloride channel [i
577 31 55.4 801 2 A37353 membrane protein 4
578 31 55.4 802 2 T45642 FtsH metalloprotei
579 31 55.4 812 2 AE0680 probable dimethyl
580 31 55.4 815 1 PHECGG phosphorylase (EC
581 31 55.4 815 2 B86009 glycogen phosphory
582 31 55.4 815 2 AB0996 glycogen phosphory
583 31 55.4 815 2 AB0996 glycogen phosphory
584 31 55.4 825 2 A48537 starch branching e
585 31 55.4 836 2 T33222 hypothetical prote
586 31 55.4 838 2 A32262 fatty-acid synthas
587 31 55.4 843 2 D96495 unknown protein [i
588 31 55.4 853 2 AB2020 hypothetical prote
589 31 55.4 885 2 G91080 hypothetical prote
590 31 55.4 886 1 JC5085 replication licens
591 31 55.4 899 2 H85925 hypothetical prote
592 31 55.4 929 2 G72677 hypothetical prote
593 31 55.4 934 2 E64235 hypothetical prote
594 31 55.4 952 2 S32954 hypothetical prote
595 31 55.4 952 2 E86147 Tln6.4 protein - A
596 31 55.4 971 2 T24866 hypothetical prote
597 31 55.4 988 2 H71338 conserved hypothet
598 31 55.4 1005 2 S73711 probable lipoprote
599 31 55.4 1016 2 AB0946 formate dehydrogen
600 31 55.4 1041 2 T31437 integrin alpha cha
601 31 55.4 1045 2 S23570 pol polyprotein ho
602 31 55.4 1140 2 D88690 protein F4IH0.3 [i
603 31 55.4 1171 1 QQKBFP pyruvate (flavodox
604 31 55.4 1202 2 S55553 LAR-interacting pr
605 31 55.4 1220 2 AD0125 exodeoxyribonuclea
606 31 55.4 1288 2 JE0363 mitogen-activated
607 31 55.4 1298 1 EDBE75 immediate-early pr
608 31 55.4 1411 2 S55123 hypothetical prote
609 31 55.4 1429 2 S06434 homeotic protein 1
610 31 55.4 1500 2 AF2027 hypothetical prote
611 31 55.4 1576 2 T28655 receptor-like hist
612 31 55.4 1576 2 T30898 receptor-like hist
613 31 55.4 1680 2 T01367 hypothetical prote

614 31 55.4 1793 2 T47897 guanine nucleotide
615 31 55.4 1972 2 S68176 TOG protein - huma
616 31 55.4 2156 1 RRVUNE genome polypotein
617 31 55.4 3036 2 T18995 hypothetical prote
618 31 55.4 3172 2 S22012 erythronolide synt
619 31 55.4 3178 2 S13595 6-deoxyerythronoli
620 31 55.4 4725 1 A44357 dynein heavy chain
621 30.5 54.5 427 1 VHVNPV nucleoprotein - pi
622 30.5 54.5 529 2 AI0986 probable membrane
623 30.5 54.5 748 2 I48744 semaphorin A - mou
624 30.5 54.5 943 2 C82559 isoleucyl-tRNA syn
625 30.5 54.5 6420 2 T30283 polyketide synthas
626 30 53.6 65 2 S59524 triose-phosphate i
627 30 53.6 69 2 E82606 transcription repr
628 30 53.6 109 2 AB3026 conserved hypothet
629 30 53.6 113 2 S26277 T cell receptor be
630 30 53.6 113 2 I38318 T-cell receptor be
631 30 53.6 113 2 S26276 hypothetical prote
632 30 53.6 114 2 E71171 hypothetical prote
633 30 53.6 116 2 T06099 hypothetical prote
634 30 53.6 119 2 G98258 hypothetical prote
635 30 53.6 119 2 C89582 protein C18A11.1 [i
636 30 53.6 123 2 C35392 hypothetical prote
637 30 53.6 123 2 T43588 hypothetical prote
638 30 53.6 123 2 S15322 hypothetical prote
639 30 53.6 142 2 AC0928 probable regulator
640 30 53.6 144 2 S38391 T-cell receptor be
641 30 53.6 145 2 S07957 hypothetical prote
642 30 53.6 148 2 H84491 hypothetical prote
643 30 53.6 150 2 T41638 probable protein k
644 30 53.6 151 2 AB0313 conserved hypothet
645 30 53.6 154 2 F64026 hypothetical prote
646 30 53.6 158 2 A85596 probable sensory t
647 30 53.6 158 2 D72305 hypothetical prote
648 30 53.6 165 2 G97231 diverged Metallo-d
649 30 53.6 172 1 NKVGD2 core protein p19 -
650 30 53.6 174 2 JCI153 hypothetical 19.7K
651 30 53.6 175 2 A95204 conserved hypothet
652 30 53.6 177 2 C69137 hypothetical prote
653 30 53.6 180 2 T09063 hypothetical prote
654 30 53.6 185 2 D82418 MutT/nudix family
655 30 53.6 188 2 C82863 recombinase XFa001
656 30 53.6 194 2 C87507 conserved hypothet
657 30 53.6 195 2 AI2549 hypothetical prote
658 30 53.6 196 2 A98071 conserved hypothet
659 30 53.6 197 2 AE2063 conserved hypothet
660 30 53.6 201 2 G59096 hypothetical prote
661 30 53.6 202 2 S56267 probable membrane
662 30 53.6 206 1 I40173 rotate phosphorib
663 30 53.6 206 2 S38626 glutathione transf
664 30 53.6 206 2 AD1898 urease accessory p
665 30 53.6 207 2 C36961 hypothetical prote
666 30 53.6 207 2 AD2350 hypothetical prote
667 30 53.6 208 2 E86014 probable enzyme yh
668 30 53.6 208 2 E91168 probable enzyme [i
669 30 53.6 208 2 S47687 hypothetical prote
670 30 53.6 209 1 XURTMC methylated-DNA-[pr
671 30 53.6 211 2 AB3178 glutathione S-tran
672 30 53.6 211 2 C82969 sarcosine oxidase
673 30 53.6 214 2 G87680 conserved hypothet
674 30 53.6 214 2 T22892 hypothetical prote
675 30 53.6 214 2 T22896 hypothetical prote
676 30 53.6 215 2 T22895 hypothetical prote
677 30 53.6 215 2 E96533 hypothetical prote
678 30 53.6 221 2 F81876 hypothetical prote
679 30 53.6 221 2 A82190 hypothetical prote
680 30 53.6 224 2 H70717 hypothetical prote
681 30 53.6 224 2 T36146 hypothetical prote
682 30 53.6 224 2 B84059 hypothetical prote
683 30 53.6 227 2 T42015 probable RNA polym
684 30 53.6 228 2 D64479 hypothetical prote
685 30 53.6 230 2 T32999 hypothetical prote
686 30 53.6 230 2 H86758 hypothetical prote

833	30	53.6	508	2	T07985	photosystem II rea	906	30	53.6	813	2	G81694	phosphorylase (EC
834	30	53.6	508	2	T07277	photosystem II chl	907	30	53.6	814	2	A71540	phosphorylase (EC
835	30	53.6	508	2	T07508	photosystem II chl	908	30	53.6	814	2	T02041	1,4-alpha-glucan b
836	30	53.6	508	2	S12129	photosystem II chl	909	30	53.6	821	2	D64119	glycogen phosphory
837	30	53.6	509	2	JA0148	photosystem II CP4	910	30	53.6	823	2	T06574	probable 1,4-alpha
838	30	53.6	509	2	AB1824	photosystem II p68	911	30	53.6	861	2	T36381	probable large ATP
839	30	53.6	509	2	T11996	Photosystem II	912	30	53.6	862	2	T36380	probable large ATP
840	30	53.6	509	2	S73243	photosystem II chl	913	30	53.6	862	2	T29133	hypothetical prote
841	30	53.6	509	2	S78322	photosystem II pro	914	30	53.6	863	2	B97504	RNA polymerase [im
842	30	53.6	509	2	T06855	photosystem II chl	915	30	53.6	863	2	AE2722	DNA-directed RNA p
843	30	53.6	513	1	S50216	translation initia	916	30	53.6	893	2	AH2007	toxin secretion AB
844	30	53.6	514	1	QJWMPB	photosystem II chl	917	30	53.6	897	2	T21688	hypothetical prote
845	30	53.6	514	2	E86738	photosystem II chl	918	30	53.6	906	2	G96621	probable disease r
846	30	53.6	514	2	G75267	multidrug resistan	919	30	53.6	908	2	B69435	signal-transducing
847	30	53.6	515	2	G70941	ABC transporter, p	920	30	53.6	914	2	S18942	hypothetical prote
848	30	53.6	522	2	T44369	hypothetical prote	921	30	53.6	945	2	F84519	probable retroelem
849	30	53.6	525	2	AF3274	pyruvate, water di	922	30	53.6	955	2	S52959	male-specific leth
850	30	53.6	534	2	A29776	hypothetical prote	923	30	53.6	973	2	AB2340	hypothetical prote
851	30	53.6	535	1	WMBEW1	glucan 1,4-alpha-g	924	30	53.6	984	2	T48216	hypothetical prote
852	30	53.6	535	2	C24187	UL21 protein - hum	925	30	53.6	1002	2	A36691	hypothetical prote
853	30	53.6	535	2	S76564	hypothetical prote	926	30	53.6	1008	2	S38003	hypothetical prote
854	30	53.6	538	2	S76175	hypothetical prote	927	30	53.6	1009	2	S44621	translation elonga
855	30	53.6	544	2	T38469	hypothetical prote	928	30	53.6	1010	2	T36383	C50C3.2 protein -
856	30	53.6	551	2	D69282	conserved hypotet	929	30	53.6	1014	2	JE0333	probable large ATP
857	30	53.6	553	2	D83640	glutamyl-tRNA synt	930	30	53.6	1021	2	T00361	klotho protein - r
858	30	53.6	563	2	T20192	hypothetical prote	931	30	53.6	1050	2	JW0092	hypothetical prote
859	30	53.6	566	2	T09154	hypothetical prote	932	30	53.6	1065	2	H95321	serine-threonine k
860	30	53.6	566	2	AC0559	hypothetical prote	933	30	53.6	1075	2	D70568	NolG efflux transp
861	30	53.6	566	2	C90691	glucose-6-phosphat	934	30	53.6	1121	2	G64103	hypothetical prote
862	30	53.6	566	2	T46219	probable solute-bi	935	30	53.6	1139	2	T23018	exodeoxyribonuclea
863	30	53.6	566	2	G85541	hypothetical prote	936	30	53.6	1150	2	T40395	hypothetical prote
864	30	53.6	566	2	E64774	hypothetical prote	937	30	53.6	1191	2	AF2501	LIM domain protein
865	30	53.6	568	2	S57830	probable membrane	938	30	53.6	1197	2	D82696	WD-repeat protein
866	30	53.6	568	2	E90364	glucose-6-phosphat	939	30	53.6	1200	2	T35599	hypothetical prote
867	30	53.6	569	2	S23542	hypothetical prote	940	30	53.6	1220	2	DJBEC3	probable DNA methy
868	30	53.6	569	2	S41806	glucose-6-phosphat	941	30	53.6	1220	2	T42573	DNA-directed DNA p
869	30	53.6	569	2	S57831	glucose-6-phosphat	942	30	53.6	1252	2	D71810	DNA-directed DNA p
870	30	53.6	571	2	T48737	probable histone a	943	30	53.6	1271	2	T43269	probable type II D
871	30	53.6	576	2	H70961	hypothetical prote	944	30	53.6	1279	2	E64709	microcystin synthe
872	30	53.6	583	2	AI3098	hypothetical prote	945	30	53.6	1435	1	BVBYL1	type IIS restricti
873	30	53.6	589	2	H84072	hypothetical prote	946	30	53.6	1558	2	T29253	guanine nucleotide
874	30	53.6	593	2	S45281	sodium/sulfate sym	947	30	53.6	1583	2	S59644	hypothetical prote
875	30	53.6	606	2	C84390	coagulation factor	948	30	53.6	1650	2	S28721	sister chromatid c
876	30	53.6	606	2	H98187	sulfate transport	949	30	53.6	1650	2	I61776	hypothetical prote
877	30	53.6	623	2	D71435	prsd protein (U891	950	30	53.6	1985	2	S27802	Munc13-2 - rat
878	30	53.6	634	2	T00054	hypothetical prote	951	30	53.6	2150	1	S27802	zinc finger protei
879	30	53.6	638	2	B83890	hypothetical prote	952	30	53.6	2150	2	T19450	hypothetical prote
880	30	53.6	639	2	JQ0607	glucan 1,4-alpha-g	953	30	53.6	2157	2	AI3009	peptide synthetase
881	30	53.6	640	1	ALASGR	glucan 1,4-alpha-g	954	30	53.6	2359	2	B96832	hypothetical prote
882	30	53.6	640	2	A29166	glucan 1,4-alpha-g	955	30	53.6	2420	2	A84652	hypothetical prote
883	30	53.6	645	2	T29818	hypothetical prote	956	30	53.6	2476	2	T34022	zonadhesin - pig
884	30	53.6	648	2	T43337	polo-like kinase-1	957	30	53.6	2566	2	E98274	hypothetical prote
885	30	53.6	652	2	T16582	hypothetical prote	958	30	53.6	3856	2	T51174	ataxia-telangiecta
886	30	53.6	653	2	C70742	probable rsbu prot	959	29.5	52.7	193	2	D83417	hypothetical prote
887	30	53.6	656	2	AC0573	outer membrane est	960	29.5	52.7	244	2	A87633	probable transcrip
888	30	53.6	676	2	H96970	endo-arabinase rel	961	29.5	52.7	300	2	S72855	hydrolase, alpha/b
889	30	53.6	683	2	B71325	conserved hypotet	962	29.5	52.7	315	2	A99192	methionine synthas
890	30	53.6	686	2	T08919	hypothetical prote	963	29.5	52.7	315	2	AI3094	nirv precursor (AF
891	30	53.6	693	2	AE0005	ATP-dependent DNA	964	29.5	52.7	584	2	D84264	nitrite reductase,
892	30	53.6	696	2	T02832	long chain fatty a	965	29.5	52.7	685	2	JC6331	hypothetical prote
893	30	53.6	702	2	T04424	probable calmoduli	966	29.5	52.7	796	2	D97065	rho-type guanine e
894	30	53.6	716	2	G44490	retrovirus-related	967	29.5	52.7	1080	2	T00587	NADH dehydrogenase
895	30	53.6	716	2	T51354	cyclic nucleotide-	968	29.5	52.7	1192	2	G70513	transketolase [imp
896	30	53.6	727	2	T08920	hypothetical prote	969	29.5	52.7	1206	2	E87072	probable ubiquitin
897	30	53.6	729	2	T06797	hypothetical prote	970	29.5	52.7	42	2	G86487	5-methyltetrahydro
898	30	53.6	740	2	AD0485	probable exported	971	29	51.8	42	2	T30614	hypothetical prote
899	30	53.6	762	2	H87466	beta-D-glucosidase	972	29	51.8	63	2	G86487	hypothetical prote
900	30	53.6	780	2	D75361	phosphoenolpyruvat	973	29	51.8	81	2	B81133	hypothetical prote
901	30	53.6	781	2	T09358	hypothetical prote	974	29	51.8	85	2	S76843	hypothetical prote
902	30	53.6	785	2	F69099	sensory transducti	975	29	51.8	87	2	S42171	hypothetical prote
903	30	53.6	799	2	T02981	1,4-alpha-glucan b	976	29	51.8	87	2	T10462	cytochrome-c oxida
904	30	53.6	799	2	T01663	1,4-alpha-glucan b	977	29	51.8	92	2	PQ0632	hypothetical prote
905	30	53.6	804	2	H75125	hypothetical prote	978	29	51.8	92	2	PQ0631	coat protein - Rem
													coat protein - lil

phosphorylase (EC
phosphorylase (EC
1,4-alpha-glucan b
glycogen phosphory
probable 1,4-alpha
probable large ATP
probable large ATP
hypothetical prote
RNA polymerase [im
DNA-directed RNA p
toxin secretion AB
hypothetical prote
probable disease r
signal-transducing
hypothetical prote
probable retroelem
male-specific leth
hypothetical prote
hypothetical prote
Ca2+-transporting
translation elonga
C50C3.2 protein -
probable large ATP
klotho protein - r
hypothetical prote
serine-threonine k
NolG efflux transp
hypothetical prote
exodeoxyribonuclea
hypothetical prote
LIM domain protein
WD-repeat protein
hypothetical prote
probable DNA methy
DNA-directed DNA p
DNA-directed DNA p
probable type II D
microcystin synthe
type IIS restricti
guanine nucleotide
hypothetical prote
sister chromatid c
hypothetical prote
Munc13-2 - rat
zinc finger protei
hypothetical prote
peptide synthetase
hypothetical prote
hypothetical prote
zonadhesin - pig
hypothetical prote
ataxia-telangiecta
hypothetical prote
probable transcrip
hydrolase, alpha/b
methionine synthas
nirv precursor (AF
nitrite reductase,
hypothetical prote
rho-type guanine e
NADH dehydrogenase
transketolase [imp
probable ubiquitin
5-methyltetrahydro
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
cytochrome-c oxida
hypothetical prote
coat protein - Rem
coat protein - lil

979 29 51.8 92 2 PQ0628 coat protein - tul
980 29 51.8 94 1 D45390 trans-activating t
981 29 51.8 97 2 H42409 glutathione transf
982 29 51.8 104 2 E82797 conserved hypothet
983 29 51.8 108 1 RPECW trp operon repress
984 29 51.8 108 2 S45254 trp operon repress
985 29 51.8 108 2 G91297 regulator for trp
986 29 51.8 108 2 A86139 trp operon repress
987 29 51.8 109 2 S45257 trp operon repress
988 29 51.8 114 2 A95025 ribosomal protein
989 29 51.8 114 2 B97896 50S ribosomal prot
990 29 51.8 115 2 F86886 50S ribosomal prot
991 29 51.8 115 2 G95084 conserved domain p
992 29 51.8 115 2 C97952 hypothetical prote
993 29 51.8 117 2 T16313 hypothetical prote
994 29 51.8 121 2 S11919 pull protein - Kle
995 29 51.8 125 1 A46315 E4 protein - human
996 29 51.8 126 2 E84047 hypothetical prote
997 29 51.8 132 1 S15618 E4 protein - human
998 29 51.8 137 2 AE0751 probable lipoprote
999 29 51.8 137 2 E64956 yedD protein - Esc
1000 29 51.8 137 2 C85810 hypothetical prote

ALIGNMENTS

RESULT 1
T11560
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)
C;Species: simian immunodeficiency virus SIVsm
A;Variety: strain E543
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11560
R;Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.;
J. Virol. 71, 1608-1620, 1997
A;Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficie
A;Reference number: Z17285; MUID:97151152; PMID:8995688
A;Accession: T11560
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1019 <HIR>
A;Cross-references: UNIPROT:P89154; UNIPARC:UPI0000105470; EMBL:U72748; NID:g1695908; PI
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; immunodeficiency

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
||| |||
Db 184 EGPKLRQW 191

RESULT 2
AG0147
probable membrane protein YP01203 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C;Accession: AG0147
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0147
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <KUR>

A;Cross-references: UNIPROT:Q8ZGS7; UNIPARC:UPI00000DC87B; GB:AL590842; PIDN:CAC90042.1;
C;Genetics:
A;Gene: YP01203
C;Superfamily: hypothetical protein ydeD

Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLROWLA 10
||| |||
Db 66 PTLQWAA 73

RESULT 3
S74539
hypothetical protein slr0740 - Synecocystis sp. (strain PCC 6803)
C;Species: Synecocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74539
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-131 <KAN>
A;Cross-references: UNIPROT:P72684; UNIPARC:UPI00000C09C4; EMBL:D90899; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Synecocystis hypothetical protein slr0740

Query Match 69.6%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLROWLA 10
||| |||
Db 29 PQLRWLA 36

RESULT 4
S46354
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)
C;Species: simian immunodeficiency virus SIVagm
A;Variety: isolate SABD37
C;Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C;Accession: S46354
R;Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.; f
EMBO J. 13, 2935-2947, 1994
A;Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A;Reference number: S46335; MUID:94298785; PMID:8026477
A;Accession: S46354
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-217 <JIN>
A;Cross-references: UNIPARC:UPI000010A53B; EMBL:U04018; NID:g466250; PIDN:AAA21512.1; PII
A;Experimental source: isolate SABD37; sabaesus monkey
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 69.6%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|| |||

Db 86 DGPRLRQW 93

RESULT 5
E84853
hypotheical protein At2g42400 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84853
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>
A;Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI000017A02D; GB:AE002093; NID:g4567312; PI
C;Genetics:
A;Gene: At2g42400
A;Map position: 2

Query Match 69.6%; Score 39; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
|||:|:
Db 344 EGETIREWL 352

RESULT 6
S30484
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30484
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30484
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
C;Superfamily: pol polyprotein

Query Match 69.6%; Score 39; DB 2; Length 656;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||:
Db 30 DGPKLRQW 37

RESULT 7
S30483
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30483
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30483
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>

A;Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111
C;Superfamily: pol polyprotein

Query Match 69.6%; Score 39; DB 2; Length 656;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||:
Db 30 DGPKLRQW 37

RESULT 8
GNLJCA
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate CAM2/Guiri
N;Contains: endonuclease (EC 3.1.1.-.); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: B38475; JQ0974
R;Tristem, M.; Hill, F.; Karpas, A.
J. Gen. Virol. 72, 721-724, 1991
A;Title: Nucleotide sequence of a Guinea-Bissau-derived human immunodeficiency virus typ
A;Reference number: A38475; MUID:91170959; PMID:2005437
A;Accession: B38475
A;Molecule type: DNA
A;Residues: 1-1034 <TRI>
A;Cross-references: UNIPROT:P24107; UNIPARC:UPI0000131EF4
A;Note: readthrough of the terminator TGA may occur between codons ATT for 564-Ile and G
C;Comment: The cleavage sites of this polyprotein have not been determined.
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F;85-183/Product: retropepsin #status predicted <RTP>
F;109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 69.6%; Score 39; DB 1; Length 1034;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||:
Db 200 DGPRLRQW 207

RESULT 9
GNLJGG
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate GH-1)
N;Contains: endonuclease (EC 3.1.1.-.); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004
C;Accession: JS0328
R;Hasegawa, A.; Tsujimoto, H.; Maki, N.; Ishikawa, K.; Miura, T.; Fukasawa, M.; Miki, K.
AIDS Res. Hum. Retroviruses 5, 593-604, 1989
A;Title: Sequence of a distinct HIV-2 isolate from Ghana showing significant divergence
A;Reference number: JS0327; MUID:90122350; PMID:2611042
A;Accession: JS0328
A;Molecule type: DNA
A;Residues: 1-1035 <HAS>
A;Cross-references: UNIPROT:P18042; UNIPARC:UPI0000174A34
A;Note: this sequence was submitted to JIPID, October 1989
C;Comment: Cleavage sites that yield the mature proteins remain to be determined.
C;Genetics:
A;Gene: pol
A;Start codon: ACA
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase;
F;85-183/Product: retropepsin #status predicted <RTP>
F;109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 69.6%; Score 39; DB 1; Length 1035;

A;Cross-references: UNIPROT:P08572; UNIPARC:UPI0000126D42; EMBL:J04210; EMBL:X05610; GB:R;Hostikka, S.L.; Kurkinen, M.; Tryggvason, K.
FEBS Lett. 216, 281-286, 1987
A;Title: Nucleotide sequence coding for the human type IV collagen alpha-2 chain cDNA reated region.
A;Reference number: S00007; MUID:87219158; PMID:3582677
A;Accession: S00007
A;Molecule type: mRNA
A;Residues: 1254-1398,'V',1400-1712 <HOS2>
A;Cross-references: UNIPARC:UPI0000173BE6; EMBL:J04210; GB:M20753; NID:g295
A;Note: 1399-Ile was also found
R;Hostikka, S.L.; Tryggvason, K.
FEBS Lett. 224, 297-305, 1987
A;Title: Extensive structural differences between genes for the alpha(1) and alpha(2) ch
A;Reference number: S02624; MUID:88083553; PMID:2826228
A;Accession: S02624
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1347-1350;1377-1383;1426-1432;1465-1471;1529-1535;1625-1630 <HOS3>
A;Cross-references: UNIPARC:UPI0000142E66; UNIPARC:UPI0000173BE7; UNIPARC:UPI0000173BE8;
A;Note: complete nucleotide sequence not shown
R;Brazel, D.; Pollner, R.; Oberbaeumer, I.; Kuehn, K.
Eur. J. Biochem. 172, 35-42, 1988
A;Title: Human basement membrane collagen (type IV): the amino acid sequence of the alph
A;Reference number: S00246; MUID:88151998; PMID:3345760
A;Accession: S00246
A;Molecule type: mRNA
A;Residues: 1-682,'G',684-1043 <BRA>
A;Cross-references: UNIPARC:UPI0000173BEC; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PID
R;Oberbaeumer, I.
submitted to the EMBL Data Library, June 1987
A;Reference number: S17678
A;Accession: S17678
A;Molecule type: mRNA
A;Residues: 1-470,'P',472-682,'G',684-1043 <OBE>
A;Cross-references: UNIPARC:UPI000016A709; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PID
R;Poeschl, E.; Pollner, R.; Kuehn, K.
EMBO J. 7, 2687-2695, 1988
A;Title: The genes for the alphas(IV) and alpha2(IV) chains of human basement membrane c
A;Reference number: S02738; MUID:89030632; PMID:2846280
A;Accession: S16911
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <POE>
A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:X12784; GB:M36963; NID:g30072; PIDN:CAA3
R;Soininen, R.; Huotari, M.; Hostikka, S.L.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 263, 17217-17220, 1988
A;Title: The structural genes for alphas1 and alpha2 chains of human type IV collagen are
A;Reference number: A92690; MUID:89034231; PMID:3182844
A;Accession: B32117
A;Molecule type: DNA
A;Residues: 1-33 <SO1>
A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; EMBL:J05039; NID:g180759; PIDN:A
R;Soininen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 264, 13565-13571, 1989
A;Title: Structural organization of the gene for the alpha-1 chain of human type IV coll
A;Reference number: S16876; MUID:89340433; PMID:2701944
A;Accession: S16877
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <SO12>
A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; NID:g180759; PIDN:AAA53097.1; PI
A;Note: this sequence was submitted to the EMBL Data Library, October 1988
R;Siebold, B.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 168, 569-575, 1987
A;Title: Construction of a model for the aggregation and cross-linking region (7S domain
is region.
A;Reference number: S00165; MUID:88029476; PMID:3117548
A;Accession: S00165
A;Molecule type: protein
A;Residues: 37-247 <SIE1>
A;Cross-references: UNIPARC:UPI0000173BED
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-Gly

R;Eble, J.A.; Golbik, R.; Mann, K.; Kuehn, K.
EMBO J. 12, 4795-4802, 1993
A;Title: The alpha-1-beta-1 integrin recognition site of the basement membrane collagen r
A;Reference number: S39614; MUID:94038963; PMID:8223488
A;Accession: S39615
A;Molecule type: protein
A;Residues: 407-570 <EBL>
A;Cross-references: UNIPARC:UPI0000173BEE
R;MacWright, R.S.; Benson, V.A.; Lovello, K.T.; van der Rest, M.; Fietzek, P.P.
Biochemistry 22, 4940-4948, 1983
A;Title: Isolation and characterization of pepsin-solubilized human basement membrane (t
A;Reference number: S16910; MUID:84053346; PMID:6416291
A;Accession: S16912
A;Molecule type: protein
A;Residues: 490-492,'X',494-496;675-677,'G',679-680,'G',682,684-685,'P' <MAC>
A;Cross-references: UNIPARC:UPI0000173BEF; UNIPARC:UPI0000173BF0
A;Experimental source: placenta
R;Glanville, R.W.; Rauter, A.
Hoppe-Seyler's Z. Physiol. Chem. 362, 943-951, 1981
A;Title: Pepsin fragments of human placental basement-membrane collagens showing interr
A;Reference number: S16908; MUID:82005835; PMID:6792033
A;Accession: B58517
A;Molecule type: protein
A;Residues: 490-492,'X',494-501,'P',503-507;952-957,'X',959-966,'X',968;984-986,'X',988-
81-1185 <GLA>
A;Cross-references: UNIPARC:UPI0000173BF1; UNIPARC:UPI0000173BF2; UNIPARC:UPI0000173BF3;
R;Killen, P.D.; Francomano, C.A.; Yamada, Y.; Modi, W.S.; O'Brien, S.J.
Hum. Genet. 77, 318-324, 1987
A;Title: Partial structure of the human alpha-2(IV) collagen chain and chromosomal local
A;Reference number: S01450; MUID:88085168; PMID:3692475
A;Accession: S01450
A;Molecule type: mRNA
A;Residues: 1040,'L',1042-1398,'V',1400-1418,'M',1420-1635,'V',1637-1712 <KIL>
A;Cross-references: UNIPARC:UPI0000072E68; EMBL:M24766; NID:g537328; PIDN:AAA52043.1; PI
R;Siebold, B.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 176, 617-624, 1988
A;Title: The arrangement of intra- and intermolecular disulfide bonds in the carboxyterm
A;Reference number: S02550; MUID:89005112; PMID:2844531
A;Accession: S02550
A;Molecule type: protein
A;Residues: 1480-1535;1545-1614;1617-1662,'H',1664-1700,'G',1705-1708;1710-1712 <SIE2>
A;Cross-references: UNIPARC:UPI0000173BF6; UNIPARC:UPI0000173BF7; UNIPARC:UPI0000173BF8;
A;Note: the sequence form Fig. 7 is inconsistent with that shown in Fig. 11 in having 17
R;Myers, J.C.; Howard, P.S.; Jelen, A.M.; Dion, A.S.; Macarak, E.J.
J. Biol. Chem. 262, 9231-9238, 1987
A;Title: Duplication of type IV collagen COOH-terminal repeats and species-specific exp
A;Reference number: A27114; MUID:87250571; PMID:2439508
A;Accession: B27114
A;Molecule type: mRNA
A;Residues: 1486-1574,'I',1576-1712 <MYE>
A;Cross-references: UNIPARC:UPI0000173BFB; EMBL:J02760; NID:g180425; PIDN:AAA58422.1; PI
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL4A2
A;Cross-references: GDB:119792; OMIM:120090
A;Map position: 13q34-13q34
A;Introns: 15/2; 33/3; 1347/1; 1380/1; 1429/1; 1468/1; 1532/1; 1527/3 #status incomplete
A;Note: the alpha 1(IV) and alpha 2(IV) chain genes are encoded on opposite strands with
C;Complex: Type IV collagen is a heterotrimer of two alpha 1(IV) chains (see PIR:CGHU4B)
domains (with disulfide and desmosine cross-links), dimeric associations among trimer ca
rupted helical domain (with disulfide and desmosine cross-links).
C;Function:
A;Description: structural component of basement membrane
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: basement membrane; cell binding; coiled coil; extracellular matrix; glycopro
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1712/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F;29-57/Domain: amino-terminal nonhelical, NH1 <NH1>
F;58-1485/Region: interrupted helical
F;362-364/Region: cell attachment (R-G-D) motif
F;784-786/Region: cell attachment (R-G-D) motif
F;868-870/Region: cell attachment (R-G-D) motif

F;889-891/Region: cell attachment (R-G-D) motif
F;970-972/Region: cell attachment (R-G-D) motif
F;1069-1071/Region: cell attachment (R-G-D) motif
F;1228-1230/Region: cell attachment (R-G-D) motif
F;1452-1454/Region: cell attachment (R-G-D) motif
F;1486-1712/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1495-1593/Domain: collagen IV carboxyl-terminal repeat <Crl>
F;1603-1708/Domain: collagen IV carboxyl-terminal repeat <Cr2>
F;42,47,51,53,137,483,485/Disulfide bonds: interchain #status predicted
F;57,87,90,102,165,168,225,239,242/Binding site: carbohydrate (Lys) (covalent) #status p
F;57/Modified site: 5-hydroxyllysine (Lys) #status atypical
F;63,75,96,114,120,123,132,150,159,186,189,198,201,213,216,219,496,499,955,964,1103,1115
F;87,90,102,165,168,225,239,242/Modified site: 5-hydroxyllysine (Lys) #status experimental
F;138/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;209/Modified site: 4-hydroxyproline (Pro) #status atypical
F;661-681/Disulfide bonds: #status predicted
F;1275/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1504-1590,1537-1593/Disulfide bonds: (or 1504-1593, 1537-1590) #status experimental
F;1549-1555,1658-1665/Disulfide bonds: #status experimental
F;1612-1705,1646-1708/Disulfide bonds: (or 1612-1708, 1646-1705) #status experimental

Query Match 67.9%; Score 38; DB 1; Length 1712;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
|||:|
Db 10 GPALRRWL 17

RESULT 15
A36925
transcription activator LysR-type CbbR - Xanthobacter flavus
C;Species: Xanthobacter flavus
C;Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text_change 09-Jul-2004
C;Accession: A36925; S13578; S35408
R;van den Bergh, E.R.E.; Dijkhuizen, L.; Meijer, W.G.
J. Bacteriol. 175, 6097-6104, 1993
A;Title: CbbR, a LysR-type transcriptional activator, is required for expression of the
A;Reference number: A36925; MUID:94012468; PMID:8407781
A;Accession: A36925
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <VAN>
A;Cross-references: UNIPROT:P25545; UNIPARC:UPI0000127169; EMBL:Z22705; NID:G297851; PID
R;Meijer, W.G.; Arnberg, A.C.; Enequist, H.G.; Terpstra, P.; Lidstrom, M.E.; Dijkhuizen,
Mol. Gen. Genet. 225, 320-330, 1991
A;Title: Identification and organization of carbon dioxide fixation genes in Xanthobacte
A;Reference number: S13573; MUID:91172133; PMID:1900916
A;Accession: S13578
A;Molecule type: DNA
A;Residues: 1-150 <MEI>
A;Cross-references: UNIPARC:UPI00001788AC; EMBL:X17252
C;Genetics:
A;Gene: cbbR
A;Start codon: GTG
C;Superfamily: transcription activator LysR-type
C;Keywords: DNA binding; transcription regulation

Query Match 67.0%; Score 37.5; DB 2; Length 333;
Best Local Similarity 72.7%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 EG-PTLRQWLA 10
|||:|
Db 265 EGLPVVRQWLA 275

RESULT 16
D70601
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacteri
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004

C;Accession: D70601
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqaures, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70601
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-306 <COL>
A;Cross-references: UNIPROT:O05576; UNIPARC:UPI00000CCAD8; GB:Z94752; GB:AL123456; NID:9;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: galU
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidyltransferase

Query Match 66.1%; Score 37; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
|||:|
Db 290 GPDLLRRWL 297

RESULT 17
T45453
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galU [similarity] - Mycobacteri
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 05-Oct-2004
C;Accession: T45453
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z22967
A;Accession: T45453
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-306 <JAM>
A;Cross-references: UNIPROT:Q9Z5G1; UNIPARC:UPI00000D438E; EMBL:AL035500; PIDN:CAB36696.3
A;Experimental source: cosmid L373
C;Genetics:
A;Note: galU
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidyltransferase

Query Match 66.1%; Score 37; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
|||:|
Db 290 GPDLLRRWL 297

RESULT 18
T36550
hypothetical protein SCH10.38c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36550
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21609
A;Accession: T36550
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-547 <MUR>
A;Cross-references: UNIPROT:Q9X8S7; UNIPARC:UPI00000DB04C; EMBL:AL049754; PIDN:CAB42047.3
A;Experimental source: strain A3 (2)
C;Genetics:

A;Gene: SCOEDB:SCH10.38C

Query Match 66.1%; Score 37; DB 2; Length 547;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
|:|:|:|
Db 57 PSLRRWLA 64

RESULT 19

A87309

hypothetical protein CC0482 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: A87309

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Exmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87309

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-777 <STO>

A;Cross-references: UNIPROT:Q9AAW1; UNIPARC:UPI00000C706B; GB:AE005673; NID:gl3421659; E

C;Genetics:

A;Gene: CC0482

C;Superfamily: cobalamin-independent methionine synthase

Query Match 66.1%; Score 37; DB 2; Length 777;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
|:|:|:|
Db 361 PEIRQWLA 368

RESULT 20

A71006

hypothetical protein PH1343 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004

C;Accession: A71006

R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: A71006

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-816 <KAW>

A;Cross-references: UNIPROT:O59068; UNIPARC:UPI0000063062; GB:AP000006; NID:g3236133; PI

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1343

Query Match 66.1%; Score 37; DB 2; Length 816;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
:| | | | |
Db 114 QGQTLRNWLS 123

RESULT 21

JC4163

DNA-binding protein 5E5 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004

C;Accession: JC4163; PC4040

R;Suzuki, E.; Kojima, N.; Yoshimura, K.; Uyemura, K.; Obata, K.; Akagawa, K. J. Biochem. 118, 122-128, 1995

A;Title: Cloning and sequence analysis of cDNA for a possible DNA-binding protein 5E5 in

A;Reference number: JC4163; MUID:96015159; PMID:8537300

A;Accession: JC4163

A;Molecule type: mRNA

A;Residues: 1-825 <SUZ>

A;Cross-references: UNIPROT:Q63003; UNIPARC:UPI0000124F19; DDBJ:D37934; NID:g531260; PID:

A;Experimental source: brain

A;Accession: PC4040

A;Molecule type: protein

A;Residues: 230-455 <SU2>

A;Cross-references: UNIPARC:UPI0000177392

C;Comment: This protein has an abundance of arginine, a glycine-rich region and a prolin

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

C;Keywords: nerve; phosphoprotein

F;436-443/Region: nuclear location signal

F;722-731/Region: proline cluster

F;62,258,345,360,404,413,570,613,635,752,820/Binding site: phosphate (Ser) (covalent) (b

Query Match 66.1%; Score 37; DB 2; Length 825;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 EGPT--LRQWL 9
| | | : | | | | |
Db 88 EGPSLPLRQWL 98

RESULT 22

S46347

pol polyprotein - simian immunodeficiency virus SIVagm (isolate SAB-1)

C;Species: simian immunodeficiency virus SIVagm

A;Variety: isolate SAB-1

C;Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999

C;Accession: S46347

R;Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.; EMBO J. 13, 2935-2947, 1994

A;Title: Mosaic genome structure of simian immunodeficiency virus from West African gree

A;Reference number: S46335; MUID:94298785; PMID:8026477

A;Accession: S46347

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1039 <JIN>

A;Cross-references: UNIPARC:UPI00001096DD; EMBL:U04005; NID:g466229; PIDN:AAA21505.1; PI

A;Experimental source: isolate SAB-1; sabaesus monkey

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993

A;Note: this reading frame extends between two stop codons and does not begin with a sta

C;Genetics:

A;Gene: pol

C;Superfamily: pol polyprotein

Query Match 66.1%; Score 37; DB 2; Length 1039;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
: | | | : | | |
Db 204 DGPRIQW 211

RESULT 23

S08436

pol polyprotein - human immunodeficiency virus type 2 D205 (fragment)

C;Species: human immunodeficiency virus type 2 D205, HIV-2 D205

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C;Accession: S08436

R;Dietrich, U.; Adamski, M.; Kreutz, R.; Seipp, A.; Kuehnel, H.; Ruebsamen-Waigmann, H. Nature 342, 948-950, 1989

A;Title: A highly divergent HIV-2-related isolate.
A;Reference number: S08434; MUID:90081881; PMID:2594088
A;Accession: S08436
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1058 <DIE>
A;Cross-references: UNIPROT:P15833; UNIPARC:UPI0000131EF6; EMBL:X16109
A;Note: this sequence was submitted to the EMBL Data Library, Aug-1989
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 66.1%; Score 37; DB 2; Length 1058;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:| |:| |
Db 221 DGPKIRQW 228

RESULT 24
S54293
regulator protein p122-RhoGAP - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 01-Sep-1995
C;Accession: S54293
R;Homma, Y.; Emori, Y.
EMBO J. 14, 286-291, 1995
A;Title: A dual functional signal mediator showing RhoGAP and phospholipase C-delta stim
A;Reference number: S54293; MUID:95137008; PMID:7835339
A;Accession: S54293
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1083 <HOM>
A;Cross-references: UNIPARC:UPI000017C9AD; EMBL:D31962

Query Match 66.1%; Score 37; DB 2; Length 1083;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | |
Db 916 EGPPLRLWRA 925

RESULT 25
S63748
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate 2-33-109
N;Contains: retropepsin (EC 3.4.23.16); RNA-directed DNA polymerase (EC 2.7.7.49)
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 2-33-109
C;Date: 20-Jul-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C;Accession: S63748; S63722
R;Yamaguchi, K.
submitted to the EMBL Data Library, July 1995
A;Reference number: S63731
A;Accession: S63748
A;Molecule type: DNA
A;Residues: 1-151 <YAM>
A;Cross-references: UNIPROT:Q72880; UNIPARC:UPI00000FFD38; EMBL:U31404; NID:g961586; PID
A;Experimental source: isolate 2-33-109
R;Yamaguchi, K.; Byrn, R.A.
Biochim. Biophys. Acta 1253, 136-140, 1995
A;Title: Clinical isolates of HIV-1 contain few pre-existing proteinase inhibitor resist
A;Reference number: S63703; MUID:96106422; PMID:8519793
A;Accession: S63722
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 23-121 <YAM>
A;Cross-references: UNIPARC:UPI000010EB23; EMBL:U31404
A;Experimental source: isolate 2-33-109

C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; hydrolase; immunodeficiency; nucleotidyltransferase
F;23-121/Product: retropepsin #status predicted <RTP>

Query Match 64.3%; Score 36; DB 2; Length 151;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:| |:| |
Db 138 DGPKVRQW 145

RESULT 26
T23485
hypothetical protein K08F4.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T23485
R;Hembry, C.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19746
A;Accession: T23485
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-200 <WIL>
A;Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN00022;
A;Experimental source: clone K08F4
C;Genetics:
A;Gene: CESP:K08F4.11
A;Map position: 4
A;Introns: 45/1; 76/1; 111/3
C;Superfamily: glutathione transferase

Query Match 64.3%; Score 36; DB 2; Length 200;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | |
Db 184 ETPKLKEWLA 193

RESULT 27
T37464
probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37464
R;Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.
submitted to the EMBL Data Library, June 1997
A;Description: Paraquat mediates differential gene expression in C. elegans.
A;Reference number: Z21702
A;Accession: T37464
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-207 <TAW>
A;Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AAB65419.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: GST3
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 64.3%; Score 36; DB 2; Length 207;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | | | |
Db 191 ETPKLKEWLA 200

RESULT 28
T45494
probable transposase A [imported] - Thiobacillus ferrooxidans transposon Tn5468
C;Species: Thiobacillus ferrooxidans
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45494
R;Oppon, J.C.; Sarnovsky, R.J.; Craig, N.L.; Rawlings, D.E.
J. Bacteriol. 180, 3007-3012, 1998
A;Title: A Tn7-like transposon is present in the glmUS region of the obligately chemoaut
A;Reference number: Z22992; MUID:98269023; PMID:9603897
A;Accession: T45494
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-278 <OPP>
A;Cross-references: UNIPROT:Q56276; UNIPARC:UPI00000B52DB; EMBL:AF032884; NID:g2653994;
A;Experimental source: ATCC 33020
C;Genetics:
A;Gene: tnsA
A;Mobile element: transposon Tn5468

Query Match 64.3%; Score 36; DB 2; Length 278;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWLA 10
|||||:
Db 215 GPTLRQFCA 223

RESULT 29
F86876
hypothetical protein yujA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86876
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: F86876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <STO>
A;Cross-references: UNIPROT:Q9CE34; UNIPARC:UPI00000C6B8E; GB:AE005176; PID:g12725061; F
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yujA

Query Match 64.3%; Score 36; DB 2; Length 312;
Best Local Similarity 40.0%; Pred. No. 67;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
:|||:::
Db 241 QGPLLKEWIS 250

RESULT 30
AB2578
membrane-bound lytic murein transglycosylase mlta [imported] - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C;Accession: AB2578
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyaivin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AB2578
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <KUR>
A;Cross-references: UNIPROT:Q8UJB9; UNIPARC:UPI00000D1687; GB:AE008688; PIDN:AAL41040.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: mlta
A;Map position: circular chromosome
C;Superfamily: membrane-bound lytic murein transglycosylase A homolog

Query Match 64.3%; Score 36; DB 2; Length 370;
Best Local Similarity 85.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TLRQWLA 10
|:|||||
Db 234 TIRQWLA 240

RESULT 31
A97360
outer membrane lipoprotein gna33 (ecoli_mltA homolog) (AF226403) [imported] - Agrobacter
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C;Accession: A97360
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <KUR>
A;Cross-references: UNIPROT:Q8UJB9; UNIPARC:UPI00000D1687; GB:AE007869; PIDN:AAK85834.1;
C;Genetics:
A;Gene: AGR_C_15
A;Map position: circular chromosome
C;Superfamily: membrane-bound lytic murein transglycosylase A homolog

Query Match 64.3%; Score 36; DB 2; Length 370;
Best Local Similarity 85.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TLRQWLA 10
|:|||||
Db 234 TIRQWLA 240

RESULT 32
T49321
related to heterokaryon incompatibility protein [imported] - Neurospora crassa
N;Alternate names: protein B13N20.60
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49321
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49321
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-481 <SCH>
A;Cross-references: UNIPROT:Q9P5Z3; UNIPARC:UPI000017B4F8; EMBL:AL355925; GSPDB:GN00116
A;Experimental source: BAC clone B13N20; strain OR74A
C;Genetics:
A;Gene: NCSP:B13N20.60
A;Map position: 6
A;Introns: 23/3

Query Match 64.3%; Score 36; DB 2; Length 481;
Best Local Similarity 75.0%; Pred. No. 1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
|||||:|

Db 376 GPTLDKWL 383

RESULT 33

A81958

probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C;Accession: A81958

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81958

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-530 <PAR>

A;Cross-references: UNIPROT:Q9JWE3; UNIPARC:UPI00000C497F; GB:AL162753; GB:AL157959; NID

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA0414

Query Match 64.3%; Score 36; DB 2; Length 530;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | |

Db 191 EMPVLRPWL 200

RESULT 34

E81015

ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: E81015

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: E81015

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-531 <TET>

A;Cross-references: UNIPROT:Q9JXI9; UNIPARC:UPI00000C4849; GB:AE002552; GB:AE002098; NID

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB2026

Query Match 64.3%; Score 36; DB 2; Length 531;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
| | | | |

Db 192 EMPVLRPWL 201

RESULT 35

S29955

surface glycoprotein - feline immunodeficiency virus

C;Species: feline immunodeficiency virus

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C;Accession: S29955

R;Rigby, M.A.
submitted to the EMBL Data Library, November 1992

A;Reference number: S29946

A;Accession: S29955

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-827 <RIG>

A;Cross-references: UNIPARC:UPI0000170E56; EMBL:X69496; NID:g397348; PIDN:CAA49250.1; PII

C;Superfamily: feline immunodeficiency virus env polyprotein

C;Keywords: glycoprotein

Query Match 64.3%; Score 36; DB 2; Length 827;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |

Db 197 EGPTLGW 204

RESULT 36

JQ2003

env polyprotein - feline immunodeficiency virus (strain UK2)

N;Alternate names: coat polyprotein

N;Contains: surface glycoprotein; transmembrane glycoprotein

C;Species: feline immunodeficiency virus

A;Note: host Felis silvestris catus (domestic cat)

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C;Accession: JQ2003; S29954

R;Rigby, M.A.; Holmes, E.C.; Pistello, M.; Mackay, A.; Brown, A.J.L.; Neil, J.C.
J. Gen. Virol. 74, 425-436, 1993

A;Title: Evolution of structural proteins of feline immunodeficiency virus: Molecular ep

A;Reference number: JQ2003; MUID:93187604; PMID:8383177

A;Accession: JQ2003

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-855 <RIG>

A;Cross-references: UNIPROT:Q04993; UNIPARC:UPI0000129FFD; GB:X69494; NID:g59267; PIDN:CF

C;Genetics:

A;Gene: env

C;Superfamily: feline immunodeficiency virus env polyprotein

C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei

F;1-610/Product: surface glycoprotein #status predicted <SPR>

F;95-111/Region: hydrophobic

F;151-178/Region: hydrophobic

F;611-855/Product: transmembrane glycoprotein #status predicted <TMP>

F;611-648/Region: hydrophobic

F;786-802/Domain: transmembrane #status predicted <TMN>

F;258,269,274,298,418,548,551,716,720,728,736/Binding site: carbohydrate (Asn) (covalent)

Query Match 64.3%; Score 36; DB 1; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |

Db 225 EGPTLGW 232

RESULT 37

JQ2004

env polyprotein - feline immunodeficiency virus (strain UK8)

N;Alternate names: coat polyprotein

N;Contains: surface glycoprotein; transmembrane glycoprotein

C;Species: feline immunodeficiency virus

A;Note: host Felis silvestris catus (domestic cat)

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C;Accession: JQ2004

R;Rigby, M.A.; Holmes, E.C.; Pistello, M.; Mackay, A.; Brown, A.J.L.; Neil, J.C.
J. Gen. Virol. 74, 425-436, 1993

A;Title: Evolution of structural proteins of feline immunodeficiency virus: Molecular ep

A;Reference number: JQ2003; MUID:93187604; PMID:8383177

A;Accession: JQ2004

A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-855 <RIG>
A;Cross-references: UNIPROT:Q04995; UNIPARC:UPI0000129FFE; GB:X69496
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; transmembrane prote
F;1-610/Product: surface glycoprotein #status predicted <SPR>
F;95-111/Region: hydrophobic
F;151-172/Region: hydrophobic
F;611-855/Product: transmembrane glycoprotein #status predicted <TMP>
F;611-648/Region: hydrophobic
F;786-802/Domain: transmembrane #status predicted <TMN>
F;258,269,274,298,418,548,551,716,720,728,736/Binding site: carbohydrate (Asn) (covalent

Query Match 64.3%; Score 36; DB 1; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 38
F45557
external glycoprotein, transmembrane protein(external glycoprotein, EGP, transmembrane p
C;Species: feline immunodeficiency virus
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: F45557
R;Maki, N.; Miyazawa, T.; Fukasawa, M.; Hasegawa, A.; Hayami, M.; Miki, K.; Mikami, T.
Arch. Virol. 123, 29-45, 1992
A;Title: Molecular characterization and heterogeneity of feline immunodeficiency virus i
A;Reference number: A45557; MUID:92198230; PMID:1312825
A;Accession: F45557
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-855 <MAK>
A;Cross-references: UNIPROT:Q02282; UNIPARC:UPI0000129FFC
A;Experimental source: strain TM2
A;Note: sequence extracted from NCBI backbone (NCBIN:89826, NCBIIP:89854)
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: glycoprotein; transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 39
A45394
env polyprotein precursor - feline immunodeficiency virus (strain UT-113)
N;Alternate names: coat polyprotein
N;Contains: surface glycoprotein; transmembrane glycoprotein
C;Species: feline immunodeficiency virus
A;Note: host Felis silvestris catus (domestic cat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A45394; S16030
R;Verschoor, E.J.; Hulskotte, E.G.J.; Ederveen, J.; Koolen, M.J.M.; Horzinek, M.C.; Rotb
Virology 193, 433-438, 1993
A;Title: Post-translational processing of the feline immunodeficiency virus envelope pre
A;Reference number: A45394; MUID:93174954; PMID:8382405
A;Accession: A45394
A;Molecule type: mRNA
A;Residues: 1-856 <VER>
A;Cross-references: UNIPROT:Q03804; UNIPARC:UPI000011D647; EMBL:X60725; NID:g1092; PIDN:
C;Comment: This protein lacks an N-terminal signal sequence, and one of the three intern
C;Genetics:

A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; transmembrane prote
F;1-611/Product: surface glycoprotein #status predicted <SGP>
F;95-111/Region: hydrophobic
F;151-169/Region: hydrophobic
F;612-856/Product: transmembrane glycoprotein #status predicted <TGP>
F;616-640/Region: hydrophobic
F;786-812/Domain: transmembrane #status predicted <TMN>
F;220,258,269,274,298,330,336,342,418,422,448,469,481,499,518,531,548,551,556,717,721,72
Query Match 64.3%; Score 36; DB 1; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 40
VCLJFP
env polyprotein precursor - feline immunodeficiency virus (strain Petaluma)
N;Alternate names: coat polyprotein
N;Contains: surface glycoprotein; transmembrane glycoprotein
C;Species: feline immunodeficiency virus
A;Note: host Felis silvestris catus (domestic cat)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: D33543
R;Talbot, R.L.; Sparger, E.E.; Lovelace, K.M.; Fitch, W.M.; Pedersen, N.C.; Luciw, P.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 5743-5747, 1989
A;Title: Nucleotide sequence and genomic organization of feline immunodeficiency virus.
A;Reference number: A33543; MUID:89345543; PMID:2762293
A;Accession: D33543
A;Molecule type: DNA
A;Residues: 1-856 <TAL>
A;Cross-references: UNIPROT:P16090; UNIPARC:UPI0000129FFA; GB:M25381
C;Comment: This protein lacks an N-terminal signal sequence, and one of the three intern
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; transmembrane prote
F;1-611/Product: surface glycoprotein #status predicted <SGP>
F;95-111/Region: hydrophobic
F;151-169/Region: hydrophobic
F;612-856/Product: transmembrane glycoprotein #status predicted <TGP>
F;616-640/Region: hydrophobic
F;786-812/Domain: transmembrane #status predicted <TMN>
F;220,258,269,274,298,330,336,342,418,422,448,481,499,518,531,548,551,717,721,729,737/B1

Query Match 64.3%; Score 36; DB 1; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 41
S23823
env polyprotein - feline immunodeficiency virus
C;Species: feline immunodeficiency virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S23823
R;Morikawa, S.; Lutz, H.; Bishop, D.H.L.
submitted to the EMBL Data Library, December 1990
A;Description: Identification of conserved and variable regions in the envelope glycopr
A;Reference number: S19886
A;Accession: S23823
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-856 <MOR>

A;Cross-references: UNIPROT:Q66975; UNIPARC:UPI000010A513; EMBL:X57002; NID:G59285; PIDN
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: polyprotein

Query Match 64.3%; Score 36; DB 2; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
|||||
Db 225 EGPTLGNW 232

RESULT 42
S19886
env polyprotein E - feline immunodeficiency virus
N;Alternate names: envelope polyprotein E
C;Species: feline immunodeficiency virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S19886
R;Morikawa, S.; Lutz, H.; Bishop, D.H.L.
submitted to the EMBL Data Library, December 1990
A;Description: Identification of conserved and variable regions in the envelope glycoprotein
A;Reference number: S19886
A;Accession: S19886
A;Molecule type: genomic RNA
A;Residues: 1-857 <MOR>
A;Cross-references: UNIPROT:Q66954; UNIPARC:UPI000010E436; EMBL:X57001; NID:G59279; PIDN
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: glycoprotein; polyprotein

Query Match 64.3%; Score 36; DB 2; Length 857;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
|||||
Db 224 EGPTLGNW 231

RESULT 43
AB1885
hypothetical protein alr0627 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB1885
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB1885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <KUR>
A;Cross-references: UNIPROT:Q8YZ60; UNIPARC:UPI000000CDDEB; GB:BA000019; PIDN:BAB72585.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0627

Query Match 64.3%; Score 36; DB 2; Length 907;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
|||||
Db 107 GPTLMSWL 114

RESULT 44

S12153
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S12153
R;Kuehnel, H.; Kreutz, R.; Ruebsamen-Waigmann, H.
Nucleic Acids Res. 18, 6142, 1990
A;Title: Nucleotide sequence of HIV-2(D194), an isolate from a Gambian case of 'Neuro-AID'
A;Reference number: S12152; MUID:91045094; PMID:2235509
A;Accession: S12153
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1032 <KUE>
A;Cross-references: UNIPROT:P17757; UNIPARC:UPI00001785D6; EMBL:X52223
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1990
C;Superfamily: pol polyprotein

Query Match 64.3%; Score 36; DB 2; Length 1032;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||:||
Db 197 DGPRLKQW 204

RESULT 45
GNLJG3
HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (macaque isolate)
N;Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
C;Species: simian immunodeficiency virus, SIV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
C;Accession: B28887
R;Chakrabarti, L.; Guyader, M.; Alizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.
Nature 328, 543-547, 1987
A;Title: Sequence of simian immunodeficiency virus from macaque and its relationship to
A;Reference number: A28887; MUID:87287230; PMID:3649576
A;Accession: B28887
A;Molecule type: DNA
A;Residues: 1-1056 <CHA>
A;Cross-references: UNIPARC:UPI0000131F1B; GB:Y00277; GB:M16403; NID:G61730; PIDN:CAA683E
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse
F;106-204/Product: retropepsin #status predicted <RTP>
F;130/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 64.3%; Score 36; DB 1; Length 1056;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||:||
Db 221 DGPKLKQW 228

RESULT 46
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; UNIPARC:UPI00000A945F; EMBL:AL391147

A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190

Query Match 64.3%; Score 36; DB 2; Length 1123;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
|||
Db 203 PTKRQWLS 210

RESULT 47
S28764
neurocan precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S28764
R;Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.
J. Biol. Chem. 267, 19536-19547, 1992
A;Title: Cloning and primary structure of neurocan, a developmentally regulated, aggregated
A;Reference number: S28764; MUID:92406907; PMID:1326557
A;Accession: S28764
A;Molecule type: mRNA
A;Residues: 1-1257 <RAU>
A;Cross-references: UNIPROT:P55067; UNIPARC:UPI0000131788; EMBL:M97161; NID:g205649; PID
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1257/Product: neurocan #status predicted <MAT>
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;364-366/Region: cell attachment (R-G-D) motif
F;953-984/Domain: EGF homology <EGF>
F;1029-1149/Domain: C-type lectin homology <LCH>
F;1156-1212/Domain: complement factor H repeat homology <FHD>
F;121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
F;944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

Query Match 64.3%; Score 36; DB 2; Length 1257;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
||||
Db 434 GPTLASWL 441

RESULT 48
T40245
probable transcription factor subunit - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40245
R;Borzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21916
A;Accession: T40245
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1339 <BOR>
A;Cross-references: UNIPROT:Q9UST7; UNIPARC:UPI0000069A03; EMBL:AL121815; PIDN:CAB58159.
A;Experimental source: strain 972h-; cosmid c336
C;Genetics:
A;Gene: SPDB:SPBC336.07
A;Map position: 2

Query Match 64.3%; Score 36; DB 2; Length 1339;
Best Local Similarity 85.7%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQW 8
||||
Db 334 GPTLPQW 340

RESULT 49
T45278
oligopeptide ABC transport protein bldKB [imported] - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45278
R;Nodwell J., R.; McGovern, K.; Losick, R.
submitted to the EMBL Data Library, August 1996
A;Description: An oligopeptide permease responsible for the import of an extracellular s
A;Reference number: Z22954
A;Accession: T45278
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-602 <NOD>
A;Cross-references: UNIPROT:P72407; UNIPARC:UPI00000BC8D6; EMBL:U68036; PIDN:AAB09555.1
A;Experimental source: strain M145
C;Genetics:
A;Note: bldKB
C;Function:
A;Description: involved in aerial mycelium formation
C;Keywords: oligopeptide transport

Query Match 63.4%; Score 35.5; DB 2; Length 602;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 EGPT-LRQWLA 10
:|||||:
Db 174 DGPTYLQQWLS 184

RESULT 50
A97655
hypothetical protein (U34346) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97655
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97655
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-120 <KUR>
A;Cross-references: UNIPROT:Q8UCW6; UNIPARC:UPI00000D1EEC; GB:AE007869; PIDN:AAK88194.1,
C;Genetics:
A;Gene: AGR_C_4462
A;Map position: circular chromosome

Query Match 62.5%; Score 35; DB 2; Length 120;
Best Local Similarity 71.4%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GPTLRQW 8
|||||:
Db 58 GPSLRRW 64

Search completed: May 12, 2006, 10:52:17
Job time : 17.2368 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-62
Perfect score: 56
Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	78.6	319	Q9RKM5_STRCO	Q9rkm5 streptomyce
2	43	76.8	297	Q7UQE4_RHOBA	Q7uqe4 rhodopirell
3	42	75.0	313	P90433_SIVCZ	P90433 chimpanzee
4	42	75.0	580	Q89RH2_BRAJA	Q89rh2 bradyrhizob
5	42	75.0	1019	POL_SIVS4	P12502 simian immu
6	42	75.0	1019	P89154_SIVCZ	P89154 chimpanzee
7	42	75.0	1019	Q72BR7_SIVCZ	Q7zbr7 chimpanzee
8	42	75.0	1019	Q7ZBR5_SIVCZ	Q7zbr5 chimpanzee
9	41	73.2	244	Q9R7K1_9SPHN	Q9r7k1 erythroba
10	41	73.2	245	O66272_9SPHN	O66272 erythroba
11	41	73.2	245	O82991_9SPHN	O82991 erythroba
12	41	73.2	249	O82989_9SPHN	O82989 erythroba
13	41	73.2	278	Q9XDV0_9SPHN	Q9xdv0 erythroba
14	41	73.2	302	Q742B3_MYCPA	Q742b3 mycobacteri
15	41	73.2	308	XERC_CORGL	Q8nnz9 corynebacte
16	41	73.2	728	Q617Q3_CAEBR	Q617q3 caenorhabdi
17	41	73.2	754	Q95Y82_CAEEL	Q95y82 caenorhabdi
18	41	73.2	791	Q73SJ6_MYCPA	Q73sj6 mycobacteri
19	41	73.2	815	Q4KS46_9VIRU	Q4ks46 orange-spot
20	41	73.2	863	Q9ST50_MAIZE	Q9st50 zea mays (m
21	41	73.2	941	Q8QUJ6_9VIRU	Q8quj6 infectious
22	40	71.4	53	Q5YWJ5_NOCFA	Q5ywj5 nocardia fa
23	40	71.4	129	Q8DHX7_SYNEL	Q8dhx7 synchococc
24	40	71.4	137	Q6FE11_ACIAD	Q6fel1 acinetobact
25	40	71.4	154	Q54XF9_DICDI	Q54xf9 dictyosteli
26	40	71.4	296	Q8ZGS7_YERPE	Q8zgs7 yersinia pe
27	40	71.4	296	Q66D06_YERPS	Q66d06 yersinia ps
28	40	71.4	326	P95613_RHIGA	P95613 rhizobium g
29	40	71.4	377	Q82PX5_STRAW	Q82px5 streptomyce
30	40	71.4	429	Q7S2Z7_NEUCR	Q7s2z7 neurospora
31	40	71.4	760	Q7G4N4_ORYSA	Q7g4n4 oryza sativ

32	39	69.6	131	2	P72684_SYNY3	P72684 synechocyst
33	39	69.6	133	2	Q4ZPV7_PSESY	Q4zpv7 pseudomonas
34	39	69.6	134	2	Q88615_PSESM	Q88615 pseudomonas
35	39	69.6	154	2	Q9WI01_9HIV2	Q9wi01 human immun
36	39	69.6	164	2	Q6MNC9_EDEBA	Q6mnc9 bdellovibri
37	39	69.6	168	2	Q9V492_DROME	Q9v492 drosophila
38	39	69.6	217	2	Q87115_SIVCZ	Q87115 chimpanzee
39	39	69.6	218	2	Q5RJN8_RAT	Q5rjn8 rattus norv
40	39	69.6	230	2	Q7ZKL1_9HIV1	Q7zkl1 human immun
41	39	69.6	245	2	O82987_9SPHN	O82987 erythroba
42	39	69.6	298	2	Q5YPT6_NOCFA	Q5ypt6 nocardia fa
43	39	69.6	310	2	Q9D8Z7_MOUSE	Q9d8z7 mus musculu
44	39	69.6	310	2	Q7SKI9_9HIV2	Q7ski9 human immun
45	39	69.6	338	2	Q4PIZ3_USTMA	Q4plz3 ustilago ma
46	39	69.6	340	2	Q8UN04_SIVCZ	Q8un04 chimpanzee
47	39	69.6	340	2	Q8UN03_SIVCZ	Q8un03 chimpanzee
48	39	69.6	349	2	Q7SKK9_9HIV2	Q7skk9 human immun
49	39	69.6	349	2	Q7SKK8_9HIV2	Q7skk8 human immun
50	39	69.6	349	2	Q7SKK6_9HIV2	Q7skk6 human immun
51	39	69.6	349	2	Q7SKK4_9HIV2	Q7skk4 human immun
52	39	69.6	349	2	Q7SKJ6_9HIV2	Q7skj6 human immun
53	39	69.6	349	2	Q7SKJ5_9HIV2	Q7skj5 human immun
54	39	69.6	349	2	Q7SKJ4_9HIV2	Q7skj4 human immun
55	39	69.6	349	2	Q7SKJ3_9HIV2	Q7skj3 human immun
56	39	69.6	349	2	Q7SKJ2_9HIV2	Q7skj2 human immun
57	39	69.6	349	2	Q7SKJ1_9HIV2	Q7skj1 human immun
58	39	69.6	349	2	Q7SKJ0_9HIV2	Q7skj0 human immun
59	39	69.6	349	2	Q7SKI8_9HIV2	Q7ski8 human immun
60	39	69.6	349	2	Q7SKI7_9HIV2	Q7ski7 human immun
61	39	69.6	349	2	Q7SKI6_9HIV2	Q7ski6 human immun
62	39	69.6	349	2	Q7SKI5_9HIV2	Q7ski5 human immun
63	39	69.6	349	2	Q7SKI4_9HIV2	Q7ski4 human immun
64	39	69.6	349	2	Q7SKI3_9HIV2	Q7ski3 human immun
65	39	69.6	349	2	Q7SKI2_9HIV2	Q7ski2 human immun
66	39	69.6	349	2	Q7SKI1_9HIV2	Q7ski1 human immun
67	39	69.6	351	2	Q9E3T1_9HIV2	Q9e3t1 human immun
68	39	69.6	351	2	Q9E3T2_9HIV2	Q9e3t2 human immun
69	39	69.6	351	2	Q9E3T4_9HIV2	Q9e3t4 human immun
70	39	69.6	396	2	Q90PU1_SIVCZ	Q90pu1 chimpanzee
71	39	69.6	450	2	Q9SLB9_ARATH	Q9slb9 arabidopsis
72	39	69.6	472	2	Q9WHZ9_9HIV2	Q9whz9 human immun
73	39	69.6	476	2	Q9WI00_9HIV2	Q9wi00 human immun
74	39	69.6	496	2	Q6CER7_YARLI	Q6cer7 yarrowia li
75	39	69.6	560	2	Q06347_9HIV2	Q06347 human immun
76	39	69.6	586	2	Q828W4_STRAW	Q828w4 streptomyce
77	39	69.6	693	2	Q5TNY2_ANOGA	Q5tny2 anopheles g
78	39	69.6	740	2	Q8XVE2_RALSO	Q8xve2 ralstonia s
79	39	69.6	756	2	Q4ZQD5_PSESY	Q4zqd5 pseudomonas
80	39	69.6	756	2	Q885P2_PSESM	Q885p2 pseudomonas
81	39	69.6	986	2	O57059_SIVCZ	O57059 chimpanzee
82	39	69.6	1017	2	Q6VG40_SIVCZ	Q6vg40 chimpanzee
83	39	69.6	1022	1	POL_SIVSP	P19505 simian immu
84	39	69.6	1022	2	Q90317_SIVCZ	Q90317 chimpanzee
85	39	69.6	1022	2	Q87956_SIVCZ	Q87956 chimpanzee
86	39	69.6	1022	2	Q87965_SIVCZ	Q87965 chimpanzee
87	39	69.6	1022	2	Q89620_SIVCZ	Q89620 chimpanzee
88	39	69.6	1022	2	Q88135_SIVCZ	Q88135 chimpanzee
89	39	69.6	1034	1	POL_HV2CA	P24107 human immun
90	39	69.6	1035	2	Q73194_9HIV2	Q73194 human immun
91	39	69.6	1035	2	Q6R793_9HIV2	Q6r793 human immun
92	39	69.6	1035	2	Q6R784_9HIV2	Q6r784 human immun
93	39	69.6	1036	1	POL_HV2RO	P04584 human immun
94	39	69.6	1049	1	POL_HV2G1	P18042 human immun
95	39	69.6	1055	1	POL_HV2ST	P20876 human immun
96	39	69.6	1055	2	Q9YTU0_9HIV2	Q9ytu0 human immun
97	39	69.6	1056	2	Q04097_SIVCZ	Q04097 chimpanzee
98	39	69.6	1059	2	Q76630_9HIV2	Q76630 human immun
99	39	69.6	1060	2	Q5QGQ1_SIVCZ	Q5qqq1 chimpanzee
100	39	69.6	1060	2	Q5QFT5_SIVCZ	Q5qft5 chimpanzee
101	39	69.6	1060	2	Q5QFR3_SIVCZ	Q5qfr3 chimpanzee
102	39	69.6	1060	2	Q5QFL2_SIVCZ	Q5qfl2 chimpanzee
103	39	69.6	1060	2	Q5QFJ6_SIVCZ	Q5qfj6 chimpanzee
104	39	69.6	1209	2	Q4WNF8_ASPFU	Q4wnf8 aspergillus

251	36	64.3	320	2	Q98YH8_9HIV1	Q98yh8	human	immun	324	36	64.3	391	2	Q75PR0_PHYPA	Q75pr0	physcomitre
252	36	64.3	321	2	Q8JB46_9HIV1	Q8jb46	human	immun	325	36	64.3	391	2	Q98Y28_9HIV1	Q98yz8	human
253	36	64.3	321	2	Q68Q60_9HIV1	Q68q60	human	immun	326	36	64.3	392	2	Q36575_9HIV1	Q36575	human
254	36	64.3	324	1	O51D1_HUMAN	Q8ngf3	homo	sapien	327	36	64.3	401	2	Q90IW3_9HIV1	Q90iw3	human
255	36	64.3	324	2	Q6WPY7_9HIV1	Q6wp7	human	immun	328	36	64.3	402	2	Q4TM76_9SPHN	Q4tmt6	erythro bact
256	36	64.3	326	2	Q4I273_GIBZE	Q4i273	gibberella		329	36	64.3	402	2	Q82R87_STRAW	Q82r87	streptomyce
257	36	64.3	326	2	Q5BJY7_RAT	Q5bjy7	rattus	norv	330	36	64.3	409	2	Q90MT8_9HIV1	Q90mt8	human
258	36	64.3	327	2	Q9D9S9_MOUSE	Q9d9s9	mus	musculu	331	36	64.3	417	2	Q8AEC2_9HIV1	Q8aec2	human
259	36	64.3	328	2	Q9EDI0_9HIV1	Q9edi0	human	immun	332	36	64.3	420	2	Q9P523_NEUCR	Q9p523	neurospora
260	36	64.3	328	2	Q9ECQ1_9HIV1	Q9ecq1	human	immun	333	36	64.3	421	2	Q7W1X1_BORPA	Q7wlx1	bordetella
261	36	64.3	328	2	Q9EIN1_9HIV1	Q9ein1	human	immun	334	36	64.3	421	2	Q7WQU8_BORBR	Q7wqu8	bordetella
262	36	64.3	328	2	Q9EE76_9HIV1	Q9ee76	human	immun	335	36	64.3	423	2	Q9QCY1_9HIV1	Q9qcy1	human
263	36	64.3	328	2	Q9EEW3_9HIV1	Q9ee76	human	immun	336	36	64.3	423	2	Q8AHP1_9HIV1	Q8ahp1	human
264	36	64.3	328	2	Q9ECX4_9HIV1	Q9ecx4	human	immun	337	36	64.3	423	2	Q8AH66_9HIV1	Q8ah66	human
265	36	64.3	328	2	Q9EEC5_9HIV1	Q9eec5	human	immun	338	36	64.3	425	2	Q8AH61_9HIV1	Q8ah61	human
266	36	64.3	328	2	Q9EKL1_9HIV1	Q9ekl1	human	immun	339	36	64.3	426	2	Q4VZ57_9HIV1	Q4vh57	human
267	36	64.3	328	2	Q9ECJ5_9HIV1	Q9ecj5	human	immun	340	36	64.3	432	2	Q9PXC7_9RETR	Q9pxc7	feline
268	36	64.3	328	2	Q9EL92_9HIV1	Q9el92	human	immun	341	36	64.3	434	2	Q4JK01_9HIV1	Q4jk01	human
269	36	64.3	328	2	Q9EDJ2_9HIV1	Q9edj2	human	immun	342	36	64.3	434	2	Q5D836_9HIV1	Q5d836	human
270	36	64.3	328	2	Q9EHC6_9HIV1	Q9ehc6	human	immun	343	36	64.3	437	1	Y4BF_RHISN	P55373	rhizobium s
271	36	64.3	328	2	Q9ECF3_9HIV1	Q9ecf3	human	immun	344	36	64.3	457	2	Q7ULE0_RHOBA	Q7ule0	rhodopirell
272	36	64.3	328	2	Q9EGW3_9HIV1	Q9egw3	human	immun	345	36	64.3	458	2	Q55QM6_CRYNE	Q55qm6	cryptococcu
273	36	64.3	329	2	Q904D4_9HIV1	Q904d4	human	immun	346	36	64.3	467	2	Q5Z206_NOCFA	Q5z206	nocardia fa
274	36	64.3	329	2	Q8QBR7_9HIV1	Q8qbr7	human	immun	347	36	64.3	468	2	Q88H13_PSEPK	Q88h13	pseudomonas
275	36	64.3	330	2	Q50AS4_9HIV1	Q50as4	human	immun	348	36	64.3	479	2	Q5D9T4_SCHJA	Q5d9t4	schistosoma
276	36	64.3	332	2	Q8Q9H6_9HIV1	Q8q9h6	human	immun	349	36	64.3	481	2	Q7RUZ6_NEUCR	Q7ruz6	neurospora
277	36	64.3	332	2	Q8Q9J8_9HIV1	Q8q9j8	human	immun	350	36	64.3	499	2	Q6RT41_9HIV1	Q6rt41	human
278	36	64.3	332	2	Q8Q9H4_9HIV1	Q8q9h4	human	immun	351	36	64.3	499	2	Q67DI8_9HIV1	Q67di8	human
279	36	64.3	332	2	Q8Q9H3_9HIV1	Q8q9h3	human	immun	352	36	64.3	499	2	Q67D97_9HIV1	Q67d97	human
280	36	64.3	332	2	Q8Q9I9_9HIV1	Q8q9i9	human	immun	353	36	64.3	499	2	Q67D81_9HIV1	Q67d81	human
281	36	64.3	332	2	Q8Q9I2_9HIV1	Q8q9i2	human	immun	354	36	64.3	499	2	Q67CP5_9HIV1	Q67cp5	human
282	36	64.3	332	2	Q7ZM79_9HIV1	Q7zm79	human	immun	355	36	64.3	499	2	Q67CB1_9HIV1	Q67cb1	human
283	36	64.3	332	2	Q7ZM65_9HIV1	Q7zm65	human	immun	356	36	64.3	499	2	Q67CA2_9HIV1	Q67ca2	human
284	36	64.3	332	2	Q7ZM00_9HIV1	Q7zm00	human	immun	357	36	64.3	499	2	Q5I1W4_9HIV1	Q5i1w4	human
285	36	64.3	332	2	Q7ZLZ8_9HIV1	Q7z1z8	human	immun	358	36	64.3	499	2	Q5I1T3_9HIV1	Q5i1t3	human
286	36	64.3	332	2	Q7ZLY1_9HIV1	Q7z1y1	human	immun	359	36	64.3	504	2	Q6EMC3_ECOLI	Q6emc3	escherichia
287	36	64.3	333	2	Q8JA16_9HIV1	Q8ja16	human	immun	360	36	64.3	504	2	Q9WTH8_9ZZZZ	Q9wth8	plasmid r10
288	36	64.3	335	2	Q82D06_STRAW	Q82d06	streptomyce		361	36	64.3	507	1	TRA6_PSEAE	P0a135	pseudomonas
289	36	64.3	335	2	Q8QAP6_9HIV1	Q8qap6	human	immun	362	36	64.3	507	1	TRA6_SHIFL	P0a135	shigella fl
290	36	64.3	335	2	Q8QAQ8_9HIV1	Q8qaq8	human	immun	363	36	64.3	507	2	Q7BT57_ECOLI	Q7bt57	escherichia
291	36	64.3	335	2	Q533H4_9HIV1	Q533h4	human	immun	364	36	64.3	507	2	Q79PC0_KLEPN	Q79pc0	klebsiella
292	36	64.3	336	1	ASB1_MOUSE	Q9wv74	mus	musculu	365	36	64.3	507	2	Q79CH1_PSEAE	Q79ch1	pseudomonas
293	36	64.3	340	2	Q50BB8_9HIV1	Q50bb8	human	immun	366	36	64.3	524	2	Q9IDL5_9HIV1	Q9idl5	human
294	36	64.3	341	2	Q6MPQ3_BDEBA	Q6mpq3	bdellovibri		367	36	64.3	530	2	Q5F558_NEIG1	Q5f558	neisseria g
295	36	64.3	342	2	Q80TI9_MOUSE	Q80ti9	mus	musculu	368	36	64.3	530	2	Q9JWE3_NEIMA	Q9jwe3	neisseria m
296	36	64.3	346	2	Q89M02_BRAJA	Q89m02	bradyrhizob		369	36	64.3	531	2	Q9JX19_NEIMB	Q9jxi9	neisseria m
297	36	64.3	349	2	Q90M47_9HIV1	Q90m47	human	immun	370	36	64.3	532	2	Q5GW75_XANOR	Q5gw75	xanthomonas
298	36	64.3	349	2	Q7SKL0_9HIV2	Q7skl0	human	immun	371	36	64.3	542	2	Q4S1Y0_TETNG	Q4sly0	tetraodon n
299	36	64.3	349	2	Q7SKK7_9HIV2	Q7skk7	human	immun	372	36	64.3	559	2	Q8UTW0_9HIV1	Q8utw0	human
300	36	64.3	349	2	Q7SKK5_9HIV2	Q7skk5	human	immun	373	36	64.3	562	2	Q99B10_9HIV1	Q99b10	human
301	36	64.3	349	2	Q7SKK3_9HIV2	Q7skk3	human	immun	374	36	64.3	566	2	Q4SJK6_TETNG	Q4sjk6	tetraodon n
302	36	64.3	349	2	Q7SKK2_9HIV2	Q7skk2	human	immun	375	36	64.3	571	2	Q7S508_NEUCR	Q7s508	neurospora
303	36	64.3	349	2	Q7SKK1_9HIV2	Q7skk1	human	immun	376	36	64.3	579	2	Q6WLD7_9BACT	Q6wld7	uncultured
304	36	64.3	349	2	Q7SKK0_9HIV2	Q7skk0	human	immun	377	36	64.3	586	2	Q70ER5_9HIV1	Q70er5	human
305	36	64.3	349	2	Q7SKJ9_9HIV2	Q7skj9	human	immun	378	36	64.3	598	2	Q6N4Y3_RHOPA	Q6n4y3	rhodopseudo
306	36	64.3	349	2	Q7SKJ8_9HIV2	Q7skj8	human	immun	379	36	64.3	607	2	Q9L8D4_POLCB	Q9l8d4	polyangium
307	36	64.3	349	2	Q7SKJ7_9HIV2	Q7skj7	human	immun	380	36	64.3	610	2	Q9L8D4_POLCB	Q9l8d4	polyangium
308	36	64.3	351	2	Q9E3T6_9HIV2	Q9e3t6	human	immun	381	36	64.3	627	2	Q6IV52_9RETR	Q6iv52	feline
309	36	64.3	351	2	Q88ZF3_9HIV1	Q88zf3	human	immun	382	36	64.3	630	2	Q5AVJ2_EMENI	Q5avj2	aspergillus
310	36	64.3	351	2	Q8Q9K2_9HIV1	Q8q9k2	human	immun	383	36	64.3	633	2	Q4KER3_PSEF5	Q4ker3	pseudomonas
311	36	64.3	351	2	Q9E3T5_9HIV2	Q9e3t5	human	immun	384	36	64.3	657	2	Q6JSB8_9HIV1	Q6jsb8	human
312	36	64.3	351	2	Q9E3T3_9HIV2	Q9e3t3	human	immun	385	36	64.3	657	2	Q6JSB7_9HIV1	Q6jsb7	human
313	36	64.3	354	2	Q9KZ09_STRCO	Q9kzq9	streptomyce		386	36	64.3	657	2	Q6JSB6_9HIV1	Q6jsb6	human
314	36	64.3	356	2	Q8AER6_9HIV1	Q8aea6	human	immun	387	36	64.3	657	2	Q6JSB5_9HIV1	Q6jsb5	human
315	36	64.3	360	2	Q90KA3_9HIV1	Q90ka3	human	immun	388	36	64.3	657	2	Q6JSB4_9HIV1	Q6jsb4	human
316	36	64.3	367	2	Q5KFI9_CRYNE	Q5kfi9	cryptococcu		389	36	64.3	657	2	Q6JSA7_9HIV1	Q6jsa7	human
317	36	64.3	370	2	Q8UJB9_AGR5	Q8ujb9	agrobacteri		390	36	64.3	657	2	Q6JSA6_9HIV1	Q6jsa6	human
318	36	64.3	371	2	Q9K567_MYCAU	Q9k567	mycobacteri		391	36	64.3	661	2	Q962C0_CAEEL	Q962c0	caenorhabdi
319	36	64.3	375	2	Q7VUT6_BORPE	Q7vut6	bordetella		392	36	64.3	669	2	Q9NDH7_CAEEL	Q9ndh7	caenorhabdi
320	36	64.3	375	2	Q7W3W9_BORPA	Q7w3w9	bordetella		393	36	64.3	674	2	Q61G29_CAEER	Q61g29	caenorhabdi
321	36	64.3	375	2	Q7WF99_BORBR	Q7wf99	bordetella		394	36	64.3	678	2	P87682_9RETR	P87682	feline
322	36	64.3	384	2	Q9J3Z1_9HIV1	Q9j3z1	human	immun	395	36	64.3	684	2	Q999V8_9RETR	Q999v8	feline
323	36	64.3	387	2	Q90GP0_9HIV1	Q90gp0	human	immun	396	36	64.3	713	2	Q7S1V3_NEUCR	Q7slv3	neurospora

397	36	64.3	764	2	Q4WK0 ASPFU	Q4wk0 aspergillus	36	64.3	1060	2	Q5QF82_SIVCZ	Q5qfs2 chimpanzee
398	36	64.3	799	2	Q82KF5_STRAW	Q82kf5 streptomyce	36	64.3	1060	2	Q5QF04_SIVCZ	Q5qfq4 chimpanzee
399	36	64.3	817	2	Q98KX1_RHILO	Q98kx1 rhizobium l	36	64.3	1060	2	Q5QFP4_SIVCZ	Q5qfp4 chimpanzee
400	36	64.3	830	2	Q521F6_MAGGR	Q521f6 magnaporthe	36	64.3	1060	2	Q5QFN4_SIVCZ	Q5qfn4 chimpanzee
401	36	64.3	851	2	Q8UNJ5_9RETR	Q8unj5 feline immu	36	64.3	1060	2	Q5QFM3_SIVCZ	Q5qfm3 chimpanzee
402	36	64.3	851	2	Q8UNJ6_9RETR	Q8unj6 feline immu	36	64.3	1060	2	Q5QFK8_SIVCZ	Q5qfk8 chimpanzee
403	36	64.3	852	2	Q66959_9RETR	Q66959 feline immu	36	64.3	1060	2	Q5IZC2_SIVCZ	Q5izc2 chimpanzee
404	36	64.3	853	2	Q66935_9RETR	Q66935 feline immu	36	64.3	1072	2	Q5BIM0_DROME	Q5bim0 drosophila
405	36	64.3	853	2	Q66957_9RETR	Q66957 feline immu	36	64.3	1072	2	Q9W592_DROME	Q9w592 drosophila
406	36	64.3	853	2	Q6IV53_9RETR	Q6iv53 feline immu	36	64.3	1073	1	POL HV2D1	P17757 human immun
407	36	64.3	854	1	ENV_FIVSD	P19030 feline immu	36	64.3	1090	2	Q4XZK5_PLACH	Q4xzk5 plasmodium
408	36	64.3	854	1	ENV_FIVWO	Q05312 feline immu	36	64.3	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis
409	36	64.3	854	2	Q90QK4_9RETR	Q90qk4 feline immu	36	64.3	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis
410	36	64.3	855	1	ENV_FIVT2	Q02282 feline immu	36	64.3	1142	1	POL HV2BB	P18096 human immun
411	36	64.3	855	1	ENV_FIVU2	Q04993 feline immu	36	64.3	1257	1	CSPG3 RAT	P55067 rattus norv
412	36	64.3	855	1	ENV_FIVU8	Q04995 feline immu	36	64.3	1330	2	Q4I254_GIBZE	Q4i254 gibberella
413	36	64.3	855	2	Q66958_9RETR	Q66958 feline immu	36	64.3	1334	2	Q5XHG9_XENLA	Q5xhg9 xenopus lae
414	36	64.3	855	2	Q66960_9RETR	Q66960 feline immu	36	64.3	1335	1	RPTOR_HUMAN	Q8n122 homo sapien
415	36	64.3	855	2	Q66961_9RETR	Q66961 feline immu	36	64.3	1335	1	RPTOR_MOUSE	Q8k4q0 mus musculu
416	36	64.3	855	2	Q66962_9RETR	Q66962 feline immu	36	64.3	1339	2	Q9UST7_SCHPO	Q9ust7 schizosacch
417	36	64.3	855	2	Q6IV18_9RETR	Q6iv18 feline immu	36	64.3	1365	2	Q4SD20_TETNG	Q4sd20 tetraodon n
418	36	64.3	855	2	Q89617_9RETR	Q89617 feline immu	36	64.3	1431	2	Q6BQ71_DEBHA	Q6bq71 debaryomyce
419	36	64.3	856	1	ENV_FIVPE	P16090 feline immu	36	64.3	1431	2	Q5B491_EMENI	Q5b491 aspergillus
420	36	64.3	856	1	ENV_FIVU1	Q03804 feline immu	36	64.3	1435	2	Q4WIA7_ASPFU	Q4wia7 aspergillus
421	36	64.3	856	2	Q03800_9RETR	Q03800 feline immu	36	64.3	1514	2	Q82BH0_STRAW	Q82bh0 streptomyce
422	36	64.3	856	2	Q03801_9RETR	Q03801 feline immu	36	64.3	3165	2	Q5AR63_EMENI	Q5ar63 aspergillus
423	36	64.3	856	2	Q66975_9RETR	Q66975 feline immu	35.5	63.4	399	2	Q7VT67_BORPE	Q7vt67 bordetella
424	36	64.3	856	2	Q9PXZ5_9RETR	Q9pxz5 feline immu	35.5	63.4	399	2	Q7W289_BORPA	Q7w289 bordetella
425	36	64.3	857	2	Q66954_9RETR	Q66954 feline immu	35.5	63.4	399	2	Q7WR56_BORBR	Q7wr56 bordetella
426	36	64.3	857	2	Q66956_9RETR	Q66956 feline immu	35.5	63.4	531	2	Q87WE7_PSESM	Q87we7 pseudomonas
427	36	64.3	859	2	Q90QK5_9RETR	Q90qk5 feline immu	35.5	63.4	600	2	Q93IU2_STRCO	Q93iu2 streptomyce
428	36	64.3	869	2	Q9VZ82_DROME	Q9vz82 drosophila	501	63.4	602	2	P72407_STRCO	P72407 streptomyce
429	36	64.3	876	2	P88142_9HIV2	P88142 human immun	502	63.4	705	2	Q6WEP8_PECCE	Q6wef8 pectobacter
430	36	64.3	898	2	Q4RLG8_TETNG	Q4rlg8 tetraodon n	503	62.5	65	2	Q984B5_RHILO	Q984b5 rhizobium l
431	36	64.3	907	2	Q8YZ60_ANASP	Q8yz60 anabaena sp	504	62.5	68	2	Q7R2G9_GIALA	Q7r2g9 giardia lam
432	36	64.3	912	2	Q9Q6Z2_9HIV1	Q9q6z2 human immun	505	62.5	72	2	Q6K895_ORYSA	Q6k895 oryza sativ
433	36	64.3	936	2	Q7ZJ29_SIVCZ	Q7zj29 simian immu	35	62.5	85	2	Q64GR5_9HIV1	Q64gr5 human immun
434	36	64.3	954	2	Q651T2_ORYSA	Q651t2 oryza sativ	35	62.5	87	2	Q65930_CAEV	Q65930 caprine art
435	36	64.3	999	2	Q8UTG9_9HIV1	Q8utg9 human immun	35	62.5	87	2	Q9WR47_9RETR	Q9wr47 brazilian c
436	36	64.3	1002	2	Q4QX95_9HIV1	Q4qx95 human immun	35	62.5	91	2	Q8Y0I5_RALSO	Q8y0i5 ralstonia s
437	36	64.3	1003	2	Q6PR23_9HIV1	Q6pr23 human immun	35	62.5	110	2	Q856U7_9CAUD	Q856u7 mycobacteri
438	36	64.3	1003	2	Q5I9G4_9HIV1	Q5i9g4 human immun	35	62.5	117	2	Q4TNK7_9SPHN	Q4tnk7 erythrobact
439	36	64.3	1005	2	Q6Y8X5_9HIV1	Q6y8x5 human immun	35	62.5	120	2	Q8UCM6_AGR75	Q8ucm6 agrobacteri
440	36	64.3	1006	2	Q5U8G6_9HIV1	Q5u8g6 human immun	35	62.5	124	2	Q5FUP7_GLUOX	Q5fup7 gluconobact
441	36	64.3	1006	2	Q596M2_9HIV1	Q596m2 human immun	35	62.5	144	2	Q7U677_SYNPK	Q7u677 synechococc
442	36	64.3	1007	2	Q66QF6_9HIV1	Q66qf6 human immun	35	62.5	148	2	Q4UK00_RICFE	Q4uk00 rickettsia
443	36	64.3	1035	1	POL_HV2KR	Q74120 human immun	35	62.5	160	2	Q73UZ7_MYCPA	Q73uz7 mycobacteri
444	36	64.3	1035	1	POL_HV2NZ	P05962 human immun	35	62.5	173	2	Q8PLV3_XANAC	Q8plv3 xanthomonas
445	36	64.3	1043	2	Q90DD2_SIVCZ	Q90dd2 chimpanzee	35	62.5	173	2	Q9B593_9HYST	Q9b593 lagidium vi
446	36	64.3	1056	1	POL_SIVM1	P05896 simian immu	35	62.5	182	2	Q94QI5_CHIBR	Q94qi5 chinchilla
447	36	64.3	1057	1	POL_SIVAI	Q02836 simian immu	35	62.5	182	2	Q9B594_9HYST	Q9b594 lagidium vi
448	36	64.3	1059	2	Q56317_9PLVG	Q56317 simian-huma	35	62.5	190	2	Q9SHP3_ARATH	Q9shp3 arabidopsis
449	36	64.3	1059	2	O11404_9PLVG	O11404 simian-huma	35	62.5	195	2	Q6N108_RHOPA	Q6n108 rhodopseudo
450	36	64.3	1059	2	Q56320_9PLVG	Q56320 simian-huma	35	62.5	197	1	BETI_PSEAE	Q88af0 pseudomonas
451	36	64.3	1059	2	Q56613_9PLVG	Q56613 simian-huma	35	62.5	197	1	BETI_PSESM	Q88af0 pseudomonas
452	36	64.3	1060	2	Q90EX3_SIVCZ	Q90ex3 simian immu	35	62.5	197	2	Q4ZM61_PSESY	Q4zm61 pseudomonas
453	36	64.3	1060	2	Q88016_SIVCZ	Q88016 chimpanzee	35	62.5	197	2	Q4K4K9_PSEF5	Q4k4k9 pseudomonas
454	36	64.3	1060	2	Q90EX8_SIVCZ	Q90ex8 simian immu	35	62.5	197	2	Q92NS0_RHIME	Q92ns0 rhizobium m
455	36	64.3	1060	2	Q8USP8_9PLVG	Q8usp8 simian-huma	35	62.5	206	2	Q8PKQ0_XANAC	Q8pkq0 xanthomonas
456	36	64.3	1060	2	Q07387_SIVCZ	Q07387 chimpanzee	35	62.5	226	2	Q7NFO5_GLOVI	Q7nfq5 gloeobacter
457	36	64.3	1060	2	Q5W9H4_SIVCZ	Q5w9h4 chimpanzee	35	62.5	232	2	Q509W0_9HIV1	Q509w0 human immun
458	36	64.3	1060	2	Q5QGP0_SIVCZ	Q5qgp0 chimpanzee	35	62.5	236	2	Q6Z4E9_ORYSA	Q6z4e9 oryza sativ
459	36	64.3	1060	2	Q5QGH9_SIVCZ	Q5qgh9 chimpanzee	35	62.5	239	2	Q7W1D9_BORPA	Q7w1d9 bordetella
460	36	64.3	1060	2	Q5QGH0_SIVCZ	Q5qgh0 chimpanzee	35	62.5	244	2	Q7W0F8_BORPE	Q7w0f8 bordetella
461	36	64.3	1060	2	Q5QGF7_SIVCZ	Q5qgf7 chimpanzee	35	62.5	244	2	Q7WP46_BORBR	Q7wp46 bordetella
462	36	64.3	1060	2	Q5QGB6_SIVCZ	Q5qgb6 chimpanzee	35	62.5	246	2	Q8N3X4_HUMAN	Q8n3x4 homo sapien
463	36	64.3	1060	2	Q5QGA9_SIVCZ	Q5qga9 chimpanzee	35	62.5	249	2	Q65SK5_MANSM	Q65sk5 manheimia
464	36	64.3	1060	2	Q5QG99_SIVCZ	Q5qg99 chimpanzee	35	62.5	257	2	Q8ELB3_OCEIH	Q8elb3 oceanobacil
465	36	64.3	1060	2	Q5QG92_SIVCZ	Q5qg92 chimpanzee	35	62.5	281	2	Q8Y2V7_RALSO	Q8y2v7 ralstonia s
466	36	64.3	1060	2	Q5QG28_SIVCZ	Q5qg28 chimpanzee	35	62.5	291	2	Q7M7W7_WOLSU	Q7m7w7 wolinnella s
467	36	64.3	1060	2	Q5QFX8_SIVCZ	Q5qfx8 chimpanzee	35	62.5	292	2	Q5GYG8_XANOR	Q5gyg8 xanthomonas
468	36	64.3	1060	2	Q5QFV7_SIVCZ	Q5qfv7 chimpanzee	35	62.5	293	2	Q4J489_AZOVI	Q4j489 azotobacter
469	36	64.3	1060	2	Q5QFU2_SIVCZ	Q5qfu2 chimpanzee	35	62.5	297	1	XERC_MYCLE	Q9cbu0 mycobacteri

543	35	62.5	305	2	Q7NRM4_CHRVO	Q7nxm4 chromobacte	616	35	62.5	577	2	Q7NL26_GLOVI	Q7nl26 gloeobacter
544	35	62.5	306	2	Q4QIH6_LEIMA	Q4qih6 leishmania	617	35	62.5	577	2	Q9IN28_9HIV1	Q9in28 human immun
545	35	62.5	307	1	PCPR_SPHCR	P52679 sphingobium	618	35	62.5	582	2	Q8AB47_BACTN	Q8ab47 bacteroides
546	35	62.5	308	2	Q5XLI6_AERHY	Q5xl16 aeromonas h	619	35	62.5	598	2	Q87Q08_VIBPA	Q87q08 vibrio para
547	35	62.5	317	2	Q8H479_ORYSA	Q8h479 oryza sativ	620	35	62.5	610	2	Q696X6_ARATH	Q696x6 arabidopsis
548	35	62.5	323	2	Q6AGE3_LEIXX	Q6age3 leifsonia x	621	35	62.5	610	2	Q696X8_ARATH	Q696x8 arabidopsis
549	35	62.5	328	2	Q5V5F2_HALMA	Q5v5f2 haloarcula	622	35	62.5	610	2	Q696Z7_ARATH	Q696z7 arabidopsis
550	35	62.5	334	2	Q5QZ31_IDILO	Q5qz31 idiomarina	623	35	62.5	610	2	Q697A2_ARATH	Q697a2 arabidopsis
551	35	62.5	335	2	Q50A80_9HIV1	Q50a80 human immun	624	35	62.5	612	1	CRY2_ARATH	Q96524 arabidopsis
552	35	62.5	342	2	Q89DI9_BRAJA	Q89di9 bradyrhizob	625	35	62.5	617	2	Q84L82_ARMRU	Q84l82 armoracia r
553	35	62.5	343	2	Q4LXZ8_9BURK	Q4lxz8 burkholderi	626	35	62.5	617	2	Q84L83_ARMRU	Q84l83 armoracia r
554	35	62.5	349	2	Q5R0S2_IDILO	Q5r0s2 idiomarina	627	35	62.5	617	2	Q84L84_ARMRU	Q84l84 armoracia r
555	35	62.5	351	2	Q72AN6_DESVH	Q72an6 desulfovibr	628	35	62.5	617	2	Q84UJ4_ARMRU	Q84uj4 armoracia r
556	35	62.5	353	2	Q5AY13_EMENI	Q5ay13 aspergillus	629	35	62.5	619	2	Q4KKX0_HUMAN	Q4kkx0 homo sapien
557	35	62.5	356	2	Q90HW3_9HIV1	Q90hw3 human immun	630	35	62.5	631	2	Q5FT72_GLUOX	Q5ft72 gluconobact
558	35	62.5	359	2	Q5V5G5_HALMA	Q5v5g5 haloarcula	631	35	62.5	635	2	Q9A882_CAUCR	Q9a882 caulobacter
559	35	62.5	370	1	Q6N0D3_RHOPA	Q6n0d3 rhodopseudo	632	35	62.5	647	2	Q6FV06_CANGA	Q6fv06 candida gla
560	35	62.5	372	1	3BHS_PIG	Q9nl19 s 3 beta-hy	633	35	62.5	653	2	Q89MC3_BRAJA	Q89mc3 bradyrhizob
561	35	62.5	372	1	Q5SJW6_THET8	Q5sejw6 thermus the	634	35	62.5	663	2	Q8IYJ8_HUMAN	Q8iyj8 homo sapien
562	35	62.5	372	2	Q72KB8_THET2	Q72kb8 thermus the	635	35	62.5	669	2	Q8DKV3_SYNEL	Q8dkv3 synechococc
563	35	62.5	372	2	Q8DM04_SYNEL	Q8dm04 synechococc	636	35	62.5	679	2	Q51U94_MAGGR	Q51u94 magnaporthe
564	35	62.5	379	1	CYB_PHADO	Q35425 phascolosor	637	35	62.5	690	1	EPB42_MOUSE	P16452 mus musculu
565	35	62.5	381	1	CYB_PHADO	Q5jlu0 phascolosor	638	35	62.5	690	1	EPB42_MOUSE	P49222 mus musculu
566	35	62.5	381	2	Q5J1U0_PHADO	Q5jlu0 phascolosor	639	35	62.5	691	2	Q4VB97_HUMAN	Q4vb97 homo sapien
567	35	62.5	383	2	Q4KHX3_PSEF5	Q4khx3 pseudomonas	640	35	62.5	691	2	Q6NXZ8_MOUSE	Q6nxz8 mus musculu
568	35	62.5	395	2	Q5P6Q0_AZOSE	Q5p6q0 azocarcus sp	641	35	62.5	705	2	Q53SX5_HUMAN	Q53sx5 homo sapien
569	35	62.5	410	2	Q7WS65_ALCEU	Q7ws65 alcaligenes	642	35	62.5	710	2	Q8RY11_ARATH	Q8ry11 arabidopsis
570	35	62.5	415	2	Q52212_9ZZZZ	Q52212 plasmid pae	643	35	62.5	712	2	Q7MUI4_PORGI	Q7mui4 porphyromon
571	35	62.5	416	2	Q8FPM9_COREF	Q8fpm9 corynebacte	644	35	62.5	713	2	Q7NL87_GLOVI	Q7nl87 gloeobacter
572	35	62.5	419	2	Q5B6X9_EMENI	Q5b6x9 aspergillus	645	35	62.5	722	2	Q6D2D3_ERWCT	Q6d2d3 erwinia car
573	35	62.5	420	2	Q8Q739_9HIV1	Q8q739 human immun	646	35	62.5	746	2	Q5F5Y4_GLUOX	Q5fsy4 gluconobact
574	35	62.5	423	2	Q8JDC1_9HIV1	Q8jdc1 human immun	647	35	62.5	758	2	Q8N9H8_HUMAN	Q8n9h8 homo sapien
575	35	62.5	433	2	Q9RD01_STRCO	Q9rd01 streptomyce	648	35	62.5	760	2	Q5GZR4_XANOR	Q5gzr4 xanthomonas
576	35	62.5	434	2	Q6JJN9_9HIV1	Q6jjn9 human immun	649	35	62.5	760	2	Q4UUK0_XANCP	Q4uuk0 xanthomonas
577	35	62.5	434	2	Q6JJN2_9HIV1	Q6jjn2 human immun	650	35	62.5	760	2	Q8P998_XANCP	Q8p998 xanthomonas
578	35	62.5	434	2	Q6JJJ9_9HIV1	Q6jjj9 human immun	651	35	62.5	760	2	Q8PL05_XANAC	Q8pl05 xanthomonas
579	35	62.5	434	2	Q6H1W7_9HIV1	Q6h1w7 human immun	652	35	62.5	766	2	Q9WWZ0_PSEPU	Q9wwz0 pseudomonas
580	35	62.5	434	2	Q5G532_9HIV1	Q5g532 human immun	653	35	62.5	791	2	Q9KU26_VIBCH	Q9ku26 vibrio chol
581	35	62.5	435	2	Q51S05_MAGGR	Q51s05 magnaporthe	654	35	62.5	793	1	KIF3C_HUMAN	O14782 homo sapien
582	35	62.5	437	2	Q5AMQ1_CANAL	Q5amq1 candida alb	655	35	62.5	793	2	Q7Z2W5_HUMAN	Q7z2w5 homo sapien
583	35	62.5	443	2	Q7R3Z8_GIALA	Q7r3z8 giardia lam	656	35	62.5	793	2	Q562F7_HUMAN	Q562f7 homo sapien
584	35	62.5	446	2	Q55LH9_CRYNE	Q55lh9 cryptococcu	657	35	62.5	795	2	O53781_MYCTU	Q562f7 mycobacteri
585	35	62.5	446	2	Q5KA18_CRYNE	Q5ka18 cryptococcu	658	35	62.5	795	2	Q7U1Q0_MYCBO	Q7ulq0 mycobacteri
586	35	62.5	446	2	Q4NNS6_9DELT	Q4nns6 anaeromyxob	659	35	62.5	803	2	Q9KKW3_VIBCH	Q9kkw3 vibrio chol
587	35	62.5	451	1	Y4337_RHIME	Q33683 rhizobium m	660	35	62.5	819	2	Q5H1C4_XANOR	Q5h1c4 xanthomonas
588	35	62.5	454	2	Q8DGB6_SYNEL	Q8dgb6 synechococc	661	35	62.5	848	2	Q4RVV3_TETNG	Q4rvv3 tetraodon n
589	35	62.5	467	2	Q6NSQ8_MOUSE	Q6nsq8 mus musculu	662	35	62.5	859	2	Q7MTD8_PORGI	Q7mtd8 porphyromon
590	35	62.5	472	1	PHR_ECOLI	P00914 escherichia	663	35	62.5	904	1	DPO1_MYCBO	P0a551 mycobacteri
591	35	62.5	472	2	Q7AGL3_ECO57	Q7agl3 escherichia	664	35	62.5	904	1	DPO1_MYCBO	P0a550 mycobacteri
592	35	62.5	472	2	Q8FJU0_ECOL6	Q8fju0 escherichia	665	35	62.5	915	2	Q6P179_HUMAN	Q6p179 homo sapien
593	35	62.5	472	2	Q8XJ9D_ECO57	Q8xj9d escherichia	666	35	62.5	925	2	Q7XSIO_ORYSA	Q7xsi0 oryza sativ
594	35	62.5	475	2	Q52QU0_PRODI	Q52qj0 prochloron	667	35	62.5	931	2	Q7WY21_PSEAE	Q7wy21 pseudomonas
595	35	62.5	480	2	O22813_ARATH	O22813 arabidopsis	668	35	62.5	933	2	Q677U1_9VIRU	Q677u1 lymphocysti
596	35	62.5	486	2	Q9SGQ0_ARATH	Q9sgq0 arabidopsis	669	35	62.5	960	2	Q7Z5K1_HUMAN	Q7z5k1 homo sapien
597	35	62.5	488	2	Q93383_CABEL	Q93383 caenorhabdi	670	35	62.5	960	2	Q9HBX2_HUMAN	Q9hbx2 homo sapien
598	35	62.5	489	2	O62100_CABEL	O62100 caenorhabdi	671	35	62.5	960	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma
599	35	62.5	491	2	Q7U9L8_SYNPX	Q7u9l8 synechococc	672	35	62.5	965	2	Q7Q3I3_ANOGA	Q7q3i3 anopheles g
600	35	62.5	501	1	PHR1_SINAL	P40115 sinapis alb	673	35	62.5	1000	2	Q4H4U6_9DEIO	Q4h4u6 deinococcus
601	35	62.5	508	2	Q6GNG6_XENLA	Q6gng6 xenopus lae	674	35	62.5	1040	2	Q9T014_ARATH	Q9t014 arabidopsis
602	35	62.5	519	2	Q8A7A4_BACTN	Q8a7a4 bacteroides	675	35	62.5	1091	1	RHG07_HUMAN	Q96qb1 homo sapien
603	35	62.5	521	2	O82507_ARATH	O82507 arabidopsis	676	35	62.5	1092	1	RHG07_MOUSE	Q9r0z9 mus musculu
604	35	62.5	522	2	Q5BGN6_EMENI	Q5bgn6 aspergillus	677	35	62.5	1092	1	RHG07_MOUSE	Q9i7s4 drosophila
605	35	62.5	522	2	Q8VKI3_MYCTU	Q8vki3 mycobacteri	678	35	62.5	1105	2	Q9I7S4_DROME	Q9i7s4 neurospora
606	35	62.5	524	2	Q9ZKU3_HELPJ	Q9zku3 helicobacte	679	35	62.5	1168	2	Q4J5B2_AZOVI	Q4j5b2 azotobacter
607	35	62.5	526	2	Q4V789_XENTR	Q4v789 xenopus tro	680	35	62.5	1170	2	Q4INJ1_GIBZE	Q4inj1 gibberella
608	35	62.5	527	2	O25564_HELPY	O25564 helicobacte	681	35	62.5	1208	2	Q5B104_EMENI	Q5b104 aspergillus
609	35	62.5	534	2	Q52SM0_9HIV2	Q52sm0 human immun	682	35	62.5	1316	2	Q6KCN9_SORMA	Q6kcn9 sordaria ma
610	35	62.5	542	2	Q6APS4_DESPS	Q6aps4 desulfotale	683	35	62.5	1322	2	Q4VK27_CAEEL	Q4vk27 caenorhabdi
611	35	62.5	553	2	Q7WGH7_BORBR	Q7wgh7 bordetella	684	35	62.5	1380	2	Q22791_CAEEL	Q22791 caenorhabdi
612	35	62.5	553	2	Q7W512_BORPA	Q7w512 bordetella	685	35	62.5	1447	2	Q4FWA4_LEIMA	Q4fwa4 leishmania
613	35	62.5	553	2	Q7VVX5_BORPE	Q7vvx5 bordetella	686	35	62.5	1515	2	Q93TX2_STIAU	Q93tx2 stigmatella
614	35	62.5	558	2	Q5JMX2_ORYSA	Q5jmx2 oryza sativ	687	35	62.5	1528	2	Q7Z5R8_HUMAN	Q7z5r8 homo sapien
615	35	62.5	569	2	Q9MA84_ARATH	Q9ma84 arabidopsis	688	35	62.5	1663	2	Q88L55_PSEPK	Q88l55 pseudomonas

689	35	62.5	1722	2	Q53L67_ORYSA	Q53l67 oryza sativ	762	34	60.7	178	2	Q9B595_CHILA	Q9b595 chinchilla
690	35	62.5	2111	2	Q5AV07_EMENI	Q5av07 aspergillus	763	34	60.7	179	2	Q94Q16_CHIBR	Q94qi6 chinchilla
691	35	62.5	2229	2	Q9EWP5_STRCO	Q9ewp5 streptomyce	764	34	60.7	182	2	Q94Q13_CHILA	Q94qi3 chinchilla
692	35	62.5	3433	1	PCX_DRÖME	P18490 drosophila	765	34	60.7	182	2	Q94Q14_CHILA	Q94qi4 chinchilla
693	35	62.5	5072	2	Q6JD66_TOXGO	Q6jd66 toxoplasma	766	34	60.7	182	2	Q94Q12_CHILA	Q94qi2 chinchilla
694	35	62.5	7599	2	Q4IKE3_GIBZE	Q4ike3 gibberella	767	34	60.7	182	2	Q9B592_CHIBR	Q9b592 chinchilla
695	34.5	61.6	294	2	Q4HCZ0_9DEIO	Q4hcz0 deinococcus	768	34	60.7	182	2	Q8HH86_CHILA	Q8hh86 chinchilla
696	34	60.7	32	2	Q4VUT0_STRPY	Q4vut0 streptococc	769	34	60.7	182	2	Q4VTW0_9SAUR	Q4vtw0 gymnodactyl
697	34	60.7	38	2	Q7YGP4_9HYST	Q7ygp4 ctenomys so	770	34	60.7	183	2	Q999D3_9HIV1	Q999d3 human immun
698	34	60.7	70	1	RK28_CYAPA	P48129 cyanophora	771	34	60.7	184	2	Q91218_PSEAE	Q91218 pseudomonas
699	34	60.7	71	2	Q64GL7_9HIV1	Q64gl7 human immun	772	34	60.7	185	2	Q4VTV4_9SAUR	Q4vtv4 gymnodactyl
700	34	60.7	85	2	Q9RS35_DEIRA	Q9rs35 deinococcus	773	34	60.7	186	2	Q8HH87_CHILA	Q8hh87 chinchilla
701	34	60.7	85	2	Q64GN1_9HIV1	Q64gn1 human immun	774	34	60.7	186	2	Q8HH85_CHILA	Q8hh85 chinchilla
702	34	60.7	93	2	Q71E12_BACSU	Q71ei2 bacillus su	775	34	60.7	186	2	Q8HBP8_CHIBR	Q8hbp8 chinchilla
703	34	60.7	97	2	Q8FPC4_COREF	Q8fpc4 corynebacte	776	34	60.7	187	2	Q4VTY4_9SAUR	Q4vty4 gymnodactyl
704	34	60.7	108	2	Q8E992_SHEON	Q8e992 shewanella	777	34	60.7	187	2	Q4VTX9_9SAUR	Q4vtx9 gymnodactyl
705	34	60.7	115	2	Q98HW9_RHILO	Q98hw9 rhizobium l	778	34	60.7	190	2	Q4VTV1_9SAUR	Q4vtv1 gymnodactyl
706	34	60.7	118	2	Q6UIA9_9HIV1	Q6uia9 human immun	779	34	60.7	191	2	Q98ZG9_9HIV1	Q98zg9 human immun
707	34	60.7	127	2	Q97468_9HIV1	Q97468 human immun	780	34	60.7	192	2	Q6N4D2_RHOPA	Q6n4d2 rhodopseudo
708	34	60.7	127	2	Q97437_9HIV1	Q97437 human immun	781	34	60.7	192	2	Q9MPW6_CHILA	Q9mpw6 chinchilla
709	34	60.7	127	2	Q97443_9HIV1	Q97443 human immun	782	34	60.7	192	2	Q6X4W7_9HYST	Q6x4w7 lagidium vi
710	34	60.7	127	2	Q97480_9HIV1	Q97480 human immun	783	34	60.7	193	2	Q7NV13_CHRVO	Q7nvi3 chromobacte
711	34	60.7	127	2	Q97441_9HIV1	Q97441 human immun	784	34	60.7	193	2	Q6X4W6_9HYST	Q6x4w6 lagidium vi
712	34	60.7	127	2	Q97463_9HIV1	Q97463 human immun	785	34	60.7	194	2	Q4VTW2_9SAUR	Q4vtw2 gymnodactyl
713	34	60.7	127	2	Q97430_9HIV1	Q97430 human immun	786	34	60.7	197	2	Q4VTX0_9SAUR	Q4vtx0 gymnodactyl
714	34	60.7	127	2	Q97428_9HIV1	Q97428 human immun	787	34	60.7	198	2	Q4VTX6_9SAUR	Q4vtx6 gymnodactyl
715	34	60.7	127	2	Q97434_9HIV1	Q97434 human immun	788	34	60.7	198	2	Q4VTV6_9SAUR	Q4vtv6 gymnodactyl
716	34	60.7	127	2	Q97461_9HIV1	Q97461 human immun	789	34	60.7	201	2	Q4ZXP3_PSESY	Q4zxp3 pseudomonas
717	34	60.7	127	2	Q97459_9HIV1	Q97459 human immun	790	34	60.7	201	2	Q887U8_PSESM	Q887u8 pseudomonas
718	34	60.7	127	2	Q97439_9HIV1	Q97439 human immun	791	34	60.7	203	2	Q6N8P8_RHOPA	Q6n8p8 rhodopseudo
719	34	60.7	127	2	Q97457_9HIV1	Q97457 human immun	792	34	60.7	204	2	Q4VTY3_9SAUR	Q4vty3 gymnodactyl
720	34	60.7	127	2	Q97432_9HIV1	Q97432 human immun	793	34	60.7	206	2	Q8Y173_RALSO	Q8y173 ralstonia s
721	34	60.7	127	2	Q97455_9HIV1	Q97455 human immun	794	34	60.7	206	2	Q6Q6C0_9HIV1	Q6q6c0 human immun
722	34	60.7	133	1	RL22_AQUAE	O66436 aquifex ao	795	34	60.7	206	2	Q6Q668_9HIV1	Q6q668 human immun
723	34	60.7	139	2	Q91Z25_9HIV1	Q91z25 human immun	796	34	60.7	207	2	Q6X4W9_9HYST	Q6x4w9 lagidium pe
724	34	60.7	144	2	Q7XWC2_ORYSA	Q7xwc2 oryza sativ	797	34	60.7	207	2	Q6X4W8_9HYST	Q6x4w8 lagidium vi
725	34	60.7	155	2	Q5DSY7_9CUCU	Q5dsy7 caryedon se	798	34	60.7	207	2	Q6X4W5_9HYST	Q6x4w5 lagidium vi
726	34	60.7	158	1	YBJN_ECOLI	P75815 escherichia	799	34	60.7	207	2	Q6Q658_9HIV1	Q6q658 human immun
727	34	60.7	158	2	Q652U6_ORYSA	Q652u6 oryza sativ	800	34	60.7	207	2	Q6Q657_9HIV1	Q6q657 human immun
728	34	60.7	158	2	Q57R86_SALCH	Q57r86 salmonella	801	34	60.7	207	2	Q6Q642_9HIV1	Q6q642 human immun
729	34	60.7	158	2	Q5PGM6_SALPA	Q5pgm6 salmonella	802	34	60.7	207	2	Q6Q641_9HIV1	Q6q641 human immun
730	34	60.7	158	2	Q8Z847_SALTI	Q8z847 salmonella	803	34	60.7	207	2	Q6Q640_9HIV1	Q6q640 human immun
731	34	60.7	158	2	Q8ZQK1_SALTY	Q8zqk1 salmonella	804	34	60.7	209	2	Q9L059_STRCO	Q9l059 streptomyce
732	34	60.7	160	2	Q742N2_MYCPA	Q742n2 mycobacteri	805	34	60.7	209	2	Q92PA0_RHIME	Q92pa0 rhizobium m
733	34	60.7	160	2	Q83DU0_COXBU	Q83du0 coxiella bu	806	34	60.7	210	2	Q7XPG5_ORYSA	Q7xpg5 oryza sativ
734	34	60.7	164	2	Q9B597_CHILA	Q9b597 chinchilla	807	34	60.7	210	2	Q4M113_9BURK	Q4m113 burkholderi
735	34	60.7	166	2	Q67F75_9HIV1	Q67f75 human immun	808	34	60.7	210	2	Q4VTW6_9SAUR	Q4vtw6 gymnodactyl
736	34	60.7	166	2	Q6TIK8_9HIV1	Q6tik8 human immun	809	34	60.7	211	2	Q98YS0_9HIV1	Q98ys0 human immun
737	34	60.7	166	2	Q6TIJ0_9HIV1	Q6tijo human immun	810	34	60.7	211	2	Q9ECL6_9HIV1	Q9ec16 human immun
738	34	60.7	166	2	Q6TII8_9HIV1	Q6tii8 human immun	811	34	60.7	213	2	Q6D7L9_ERWCT	Q6d7l9 erwinia car
739	34	60.7	166	2	Q6TII7_9HIV1	Q6tii7 human immun	812	34	60.7	214	2	Q63IV1_BURPS	Q63iv1 burkholderi
740	34	60.7	166	2	Q6TIG8_9HIV1	Q6tig8 human immun	813	34	60.7	214	2	Q62ED0_BURMA	Q62ed0 burkholderi
741	34	60.7	166	2	Q6TIG7_9HIV1	Q6tig7 human immun	814	34	60.7	215	2	Q6UZR2_9HIV1	Q6uzr2 human immun
742	34	60.7	166	2	Q6TIF4_9HIV1	Q6tif4 human immun	815	34	60.7	217	2	Q4P6B1_USTMA	Q4p6b1 ustilago ma
743	34	60.7	166	2	Q6TIF2_9HIV1	Q6tif2 human immun	816	34	60.7	218	2	Q5AZJ2_EMENI	Q5azj2 aspergillus
744	34	60.7	166	2	Q6TIE1_9HIV1	Q6tie1 human immun	817	34	60.7	218	2	Q4VTV8_9SAUR	Q4vtv8 gymnodactyl
745	34	60.7	166	2	Q6TID8_9HIV1	Q6tid8 human immun	818	34	60.7	218	2	Q5MI28_9HIV1	Q5mi28 human immun
746	34	60.7	166	2	Q6TIA7_9HIV1	Q6tia7 human immun	819	34	60.7	219	2	Q89943_9HIV1	Q89943 human immun
747	34	60.7	166	2	Q6TI98_9HIV1	Q6ti98 human immun	820	34	60.7	219	2	Q9ID86_9HIV1	Q9id86 human immun
748	34	60.7	166	2	Q6TI90_9HIV1	Q6ti90 human immun	821	34	60.7	220	1	NADD_YERPE	Q8zdbg1 yersinia pe
749	34	60.7	166	2	Q6TI73_9HIV1	Q6ti73 human immun	822	34	60.7	220	2	Q66DE6_YERPS	Q66de6 yersinia ps
750	34	60.7	166	2	Q6TI63_9HIV1	Q6ti63 human immun	823	34	60.7	220	2	Q85H40_9SAUR	Q85h40 pachydactyl
751	34	60.7	166	2	Q6TI37_9HIV1	Q6ti37 human immun	824	34	60.7	220	2	Q4VTV5_9SAUR	Q4vtv5 gymnodactyl
752	34	60.7	166	2	Q6TI21_9HIV1	Q6ti21 human immun	825	34	60.7	221	2	Q957I9_LACAG	Q957i9 lacerta agi
753	34	60.7	166	2	Q6TI20_9HIV1	Q6ti20 human immun	826	34	60.7	221	2	Q8M2E4_9SAUR	Q8m2e4 pachydactyl
754	34	60.7	167	2	Q9B599_CHILA	Q9b599 chinchilla	827	34	60.7	221	2	Q8M2E1_9SAUR	Q8m2e1 pachydactyl
755	34	60.7	168	2	Q4TFP2_TETNG	Q4tfp2 tetraodon n	828	34	60.7	221	2	Q8M2D2_9SAUR	Q8m2d2 pachydactyl
756	34	60.7	169	2	Q94QI1_9HYST	Q94qi1 octomys mim	829	34	60.7	221	2	Q8M2E2_9SAUR	Q8m2e2 pachydactyl
757	34	60.7	169	2	Q9ECC2_9HIV1	Q9ecc2 human immun	830	34	60.7	221	2	Q5MI49_9HIV1	Q5mi49 human immun
758	34	60.7	169	2	Q9QF21_9HIV1	Q9qf21 human immun	831	34	60.7	222	2	Q9ID73_9HIV1	Q9id73 human immun
759	34	60.7	170	2	Q9B5A0_CHILA	Q9b5a0 chinchilla	832	34	60.7	222	2	Q5RMI0_9HIV1	Q5rmi0 human immun
760	34	60.7	171	2	Q8Q8W3_9HIV1	Q8q8w3 human immun	833	34	60.7	223	2	Q4VTW7_9SAUR	Q4vtw7 gymnodactyl
761	34	60.7	178	2	Q84XL1_LYCES	Q84xl1 lycopersico	834	34	60.7	223	2	Q4VTW4_9SAUR	Q4vtw4 gymnodactyl

835	34	60.7	223	2	Q999D2_9HIV1	Q999d2	human	immun	908	34	60.7	238	2	Q5MHW4_9HIV1	Q5mhw4	human	immun
836	34	60.7	224	2	Q98X05_9HIV1	Q98x05	human	immun	909	34	60.7	238	2	Q509Q1_9HIV1	Q509q1	human	immun
837	34	60.7	224	2	Q98Ww5_9HIV1	Q98ww5	human	immun	910	34	60.7	239	2	Q7SM78_9HIV1	Q7sm78	human	immun
838	34	60.7	224	2	Q98Wz9_9HIV1	Q98wz9	human	immun	911	34	60.7	239	2	Q66w37_9HIV1	Q66w37	human	immun
839	34	60.7	224	2	Q98X19_9HIV1	Q98x19	human	immun	912	34	60.7	240	2	Q7Zkp5_9HIV1	Q7zkp5	human	immun
840	34	60.7	224	2	Q98WZ1_9HIV1	Q98wz1	human	immun	913	34	60.7	241	2	Q90M39_9HIV1	Q90m39	human	immun
841	34	60.7	224	2	Q99EU4_9HIV1	Q99eu4	human	immun	914	34	60.7	241	2	Q5RM97_9HIV1	Q5rm97	human	immun
842	34	60.7	224	2	Q98X02_9HIV1	Q98x02	human	immun	915	34	60.7	242	2	Q9T7E0_9HYST	Q9t7e0	trinomys	gr
843	34	60.7	224	2	Q5MHX2_9HIV1	Q5mhx2	human	immun	916	34	60.7	242	2	Q9T7D9_9HYST	Q9t7d9	trinomys	gr
844	34	60.7	225	2	Q4VTv9_9SAUR	Q4vtv9	gymnodactyl		917	34	60.7	242	2	Q9T7E1_9HYST	Q9t7e1	trinomys	gr
845	34	60.7	225	2	Q999G4_9HIV1	Q999g4	human	immun	918	34	60.7	242	2	Q8SF17_SIGHI	Q8sfi7	sigmodon	hi
846	34	60.7	226	2	Q4VTW9_9SAUR	Q4vtw9	gymnodactyl		919	34	60.7	242	2	Q8SELO_9RODE	Q8sel0	sigmodon	ma
847	34	60.7	226	2	Q8Q3T7_9HIV1	Q8q3t7	human	immun	920	34	60.7	242	2	Q9T4B5_9HYST	Q9t4b5	trinomys	gr
848	34	60.7	226	2	Q99ET8_9HIV1	Q99et8	human	immun	921	34	60.7	242	2	Q4VTX4_9SAUR	Q4vtx4	gymnodactyl	
849	34	60.7	226	2	Q99ET7_9HIV1	Q99et7	human	immun	922	34	60.7	243	2	Q5MI35_9HIV1	Q5mi35	human	immun
850	34	60.7	226	2	Q99EV4_9HIV1	Q99ev4	human	immun	923	34	60.7	243	2	Q5MHT7_9HIV1	Q5mht7	human	immun
851	34	60.7	226	2	Q99D46_9HIV1	Q99d46	human	immun	924	34	60.7	244	2	Q5B7M7_EMENI	Q5b7m7	aspergillus	
852	34	60.7	226	2	Q66W22_9HIV1	Q66w22	human	immun	925	34	60.7	244	2	Q90BB3_9HIV1	Q90bb3	human	immun
853	34	60.7	227	2	Q4VTv7_9SAUR	Q4vtv7	gymnodactyl		926	34	60.7	244	2	Q7ZKJ9_9HIV1	Q7zkj9	human	immun
854	34	60.7	227	2	Q90RM9_9HIV1	Q90rm9	human	immun	927	34	60.7	244	2	Q7ZKJ4_9HIV1	Q7zkj4	human	immun
855	34	60.7	227	2	Q992D3_9HIV1	Q992d3	human	immun	928	34	60.7	244	2	Q5MI44_9HIV1	Q5mi44	human	immun
856	34	60.7	227	2	Q992B3_9HIV1	Q992b3	human	immun	929	34	60.7	244	2	Q5MI37_9HIV1	Q5mi37	human	immun
857	34	60.7	227	2	Q5MHV6_9HIV1	Q5mhv6	human	immun	930	34	60.7	244	2	Q5MI27_9HIV1	Q5mi27	human	immun
858	34	60.7	228	2	Q9MHR7_9SAUR	Q9mhr7	pachydactyl		931	34	60.7	244	2	Q5MHU6_9HIV1	Q5mhu6	human	immun
859	34	60.7	228	2	Q4VTW5_9SAUR	Q4vtw5	gymnodactyl		932	34	60.7	245	2	Q7ZKM6_9HIV1	Q7zkm6	human	immun
860	34	60.7	229	2	Q70EG3_9RODE	Q70eg3	saccostomus		933	34	60.7	245	2	Q7ZKI3_9HIV1	Q7zki3	human	immun
861	34	60.7	229	2	Q9Q302_9HIV1	Q9q302	human	immun	934	34	60.7	245	2	Q7ZF60_9HIV1	Q7zpf60	human	immun
862	34	60.7	229	2	Q9Q380_9HIV1	Q9q380	human	immun	935	34	60.7	245	2	Q6QJU1_9HIV1	Q6qju1	human	immun
863	34	60.7	229	2	Q9Q316_9HIV1	Q9q316	human	immun	936	34	60.7	245	2	Q5MI16_9HIV1	Q5mi16	human	immun
864	34	60.7	229	2	Q5MHV8_9HIV1	Q5mhv8	human	immun	937	34	60.7	246	2	Q7ZKL6_9HIV1	Q7zkl6	human	immun
865	34	60.7	229	2	Q50BA3_9HIV1	Q50ba3	human	immun	938	34	60.7	247	2	Q7ZKE0_9HIV1	Q7zke0	human	immun
866	34	60.7	229	2	Q4VTX1_9SAUR	Q4vtx1	gymnodactyl		939	34	60.7	247	2	Q5MI50_9HIV1	Q5mi50	human	immun
867	34	60.7	230	2	Q4VTU9_9SAUR	Q4vtu9	gymnodactyl		940	34	60.7	247	2	Q5MI06_9HIV1	Q5mi06	human	immun
868	34	60.7	230	2	Q9Q392_9HIV1	Q9q392	human	immun	941	34	60.7	247	2	Q5MHV0_9HIV1	Q5mhv0	human	immun
869	34	60.7	230	2	Q992E6_9HIV1	Q992e6	human	immun	942	34	60.7	247	2	Q5MHU2_9HIV1	Q5mhu2	human	immun
870	34	60.7	230	2	Q5MI09_9HIV1	Q5mi09	human	immun	943	34	60.7	249	2	Q72KG2_9HIV1	Q7zkg2	human	immun
871	34	60.7	231	2	Q4VTv3_9SAUR	Q4vtv3	gymnodactyl		944	34	60.7	249	2	Q5RSY6_9HIV1	Q5rsy6	human	immun
872	34	60.7	232	2	Q6N2F2_RHOPA	Q6n2f2	rhodopseudo		945	34	60.7	249	2	Q5RSY5_9HIV1	Q5rsy5	human	immun
873	34	60.7	232	2	Q4VTX8_9SAUR	Q4vtx8	gymnodactyl		946	34	60.7	249	2	Q5RR44_9HIV1	Q5rr44	human	immun
874	34	60.7	232	2	Q4VTW1_9SAUR	Q4vtw1	gymnodactyl		947	34	60.7	249	2	Q5RQ83_9HIV1	Q5rq83	human	immun
875	34	60.7	232	2	Q999Q8_9HIV1	Q999q8	human	immun	948	34	60.7	250	2	Q90M40_9HIV1	Q90m40	human	immun
876	34	60.7	232	2	Q85Q52_9HYST	Q85q52	thrichomys		949	34	60.7	250	2	Q8J853_9HIV1	Q8j853	human	immun
877	34	60.7	233	2	Q4VTY2_9SAUR	Q4vty2	gymnodactyl		950	34	60.7	250	2	Q90M44_9HIV1	Q90m44	human	immun
878	34	60.7	233	2	Q4VTW8_9SAUR	Q4vtw8	gymnodactyl		951	34	60.7	250	2	Q71309_9HIV1	Q71309	human	immun
879	34	60.7	233	2	Q4VTW3_9SAUR	Q4vtw3	gymnodactyl		952	34	60.7	250	2	Q991W5_9HIV1	Q991w5	human	immun
880	34	60.7	233	2	Q6QJR5_9HIV1	Q6qjr5	human	immun	953	34	60.7	250	2	Q71310_9HIV1	Q71310	human	immun
881	34	60.7	233	2	Q5MHX5_9HIV1	Q5mhx5	human	immun	954	34	60.7	250	2	Q71308_9HIV1	Q71308	human	immun
882	34	60.7	234	2	Q4K3V9_PSEF5	Q4k3v9	pseudomonas		955	34	60.7	251	2	Q6QJR6_9HIV1	Q6qjr6	human	immun
883	34	60.7	234	2	Q4VTX7_9SAUR	Q4vtx7	gymnodactyl		956	34	60.7	251	2	Q5HLW9_STAEP	Q5hlw9	staphylococ	
884	34	60.7	234	2	Q999E7_9HIV1	Q999e7	human	immun	957	34	60.7	251	2	Q8CRE9_STAEP	Q8cre9	staphylococ	
885	34	60.7	234	2	Q50B99_9HIV1	Q50b99	human	immun	958	34	60.7	251	2	Q5RNL8_9HIV1	Q5rnl8	human	immun
886	34	60.7	234	2	Q50B86_9HIV1	Q50b86	human	immun	959	34	60.7	252	2	Q8FT86_COREF	Q8ft86	corynebacte	
887	34	60.7	234	2	Q50B27_9HIV1	Q50b27	human	immun	960	34	60.7	252	2	Q6NH53_CORDI	Q6nh53	corynebacte	
888	34	60.7	234	2	Q509V2_9HIV1	Q509v2	human	immun	961	34	60.7	252	2	Q8NQ77_CORDI	Q8nq77	corynebacte	
889	34	60.7	235	2	Q4VTY1_9SAUR	Q4vty1	gymnodactyl		962	34	60.7	252	2	Q4VTV2_9SAUR	Q4vtv2	gymnodactyl	
890	34	60.7	235	2	Q4VTX5_9SAUR	Q4vtx5	gymnodactyl		963	34	60.7	253	2	Q6N745_RHOPA	Q6n745	rhodopseudo	
891	34	60.7	235	2	Q4VTV0_9SAUR	Q4vtv0	gymnodactyl		964	34	60.7	253	2	Q90CM2_9HIV1	Q90cm2	human	immun
892	34	60.7	235	2	Q5MHW8_9HIV1	Q5mhv8	human	immun	965	34	60.7	253	2	Q7ZKL3_9HIV1	Q7zkl3	human	immun
893	34	60.7	235	2	Q50B76_9HIV1	Q50b76	human	immun	966	34	60.7	254	2	Q5RR25_9HIV1	Q5rr25	human	immun
894	34	60.7	236	2	Q53A80_9HIV1	Q53a80	human	immun	967	34	60.7	255	2	Q6TV17_9HYST	Q6tv17	isothrix	bi
895	34	60.7	237	2	Q6XNT0_9SAUR	Q6xnt0	phelsuma	gu	968	34	60.7	255	2	Q7ZKK6_9HIV1	Q7zkk6	human	immun
896	34	60.7	237	2	Q7ZKD7_9HIV1	Q7zkd7	human	immun	969	34	60.7	255	2	Q7ZKF7_9HIV1	Q7zke7	human	immun
897	34	60.7	237	2	Q68Q78_9HIV1	Q68q78	human	immun	970	34	60.7	255	2	Q7ZKE7_9HIV1	Q7zke7	human	immun
898	34	60.7	238	2	Q6XNZ7_9SAUR	Q6xnz7	pachydactyl		971	34	60.7	255	2	Q7ZF24_9HIV1	Q7zsf24	human	immun
899	34	60.7	238	2	Q6XNZ3_9SAUR	Q6xnz3	phelsuma	mu	972	34	60.7	255	2	Q6QTR0_9HIV1	Q6qtr0	human	immun
900	34	60.7	238	2	Q6XNY9_9SAUR	Q6xny9	phelsuma	ma	973	34	60.7	255	2	Q6QTP2_9HIV1	Q6qtp2	human	immun
901	34	60.7	238	2	Q6XNY6_9SAUR	Q6xny6	phelsuma	li	974	34	60.7	255	2	Q6QTP1_9HIV1	Q6qtp1	human	immun
902	34	60.7	238	2	Q6XNY3_9SAUR	Q6xny3	phelsuma	qu	975	34	60.7	255	2	Q5RRL7_9HIV1	Q5rrl7	human	immun
903	34	60.7	238	2	Q6XNY2_9SAUR	Q6xny2	phelsuma	ab	976	34	60.7	255	2	Q5MHW0_9HIV1	Q5mhw0	human	immun
904	34	60.7	238	2	Q6XNY1_9SAUR	Q6xny1	phelsuma	ab	977	34	60.7	255	2	Q4TUU2_9HIV1	Q4tuu2	human	immun
905	34	60.7	238	2	Q6XNY0_9SAUR	Q6xny0	phelsuma	ab	978	34	60.7	256	2	Q4VTU8_9SAUR	Q4vtu8	gymnodactyl	
906	34	60.7	238	2	Q6XNS9_9SAUR	Q6xns9	phelsuma	gu	979	34	60.7	256	2	Q5RS93_9HIV1	Q5rs93	human	immun
907	34	60.7	238	2	Q8AFU0_9HIV1	Q8afu0	human	immun	980	34	60.7	257	2	Q5RS99_9HIV1	Q5rs99	human	immun

```
981 34 60.7 258 2 Q85Q43 9HYST Q85Q43 trichomys
982 34 60.7 258 2 Q5RT01_9HIV1 Q5rt01 human immun
983 34 60.7 259 2 Q4UP49_XANCP Q4up49 xanthomonas
984 34 60.7 259 2 Q8P3M5_XANCP Q8p3m5 xanthomonas
985 34 60.7 259 2 Q8PF17_XANAC Q8pf17 xanthomonas
986 34 60.7 259 2 Q8AFU1_9HIV1 Q8afu1 human immun
987 34 60.7 259 2 Q8AFT7_9HIV1 Q8aft7 human immun
988 34 60.7 261 2 Q8J877_9HIV1 Q8j877 human immun
989 34 60.7 261 2 Q5IR78_9HIV1 Q5ir78 human immun
990 34 60.7 262 2 Q9YQ24_SIVCZ Q9yq24 chimpanzee
991 34 60.7 263 2 Q5LVW4_SILPO Q5lvw4 silicibacte
992 34 60.7 264 2 Q4VTU1_9SAUR Q4vtul gymnodactyl
993 34 60.7 264 2 Q5VAC6_9HIV1 Q5vac6 human immun
994 34 60.7 264 2 Q5RQO0_9HIV1 Q5rqo0 human immun
995 34 60.7 264 2 Q5RMH8_9HIV1 Q5rmh8 human immun
996 34 60.7 266 2 Q34813_9HYST Q34813 isothrix bi
997 34 60.7 266 2 Q34820_9HYST Q34820 isothrix pa
998 34 60.7 266 2 Q34421_9HYST Q34421 echimys chr
999 34 60.7 266 2 Q34812_9HYST Q34812 isothrix bi
1000 34 60.7 266 2 Q34307_9HYST Q34307 dactylomys
```

ALIGNMENTS

```
RESULT 1
Q9RKM5 STRCO
ID Q9RKM5_STRCO PRELIMINARY; PRT; 319 AA.
AC Q9RKM5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative MerR family transcriptional regulator.
GN OrderedLocusNames=SCO4102; ORFNames=SCD17.06C;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939118; CAB56383.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS50937; HTH_MERR 2; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;
```

```
Query Match 78.6%; Score 44; DB 2; Length 319;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 258 DGPRLREWLA 267
:|||||:
:|||||:
```

```
RESULT 2
Q7UQE4 RHOBA
ID Q7UQE4_RHOBA PRELIMINARY; PRT; 297 AA.
AC Q7UQE4;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB6375;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294144; CAD74759.1; -; Genomic_DNA.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR003169; GYF.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50829; GYF; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475F670F02C78E9B CRC64;
```

```
Query Match 76.8%; Score 43; DB 2; Length 297;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 176 DGPTMKQWIS 185
:|||||:
:|||||:
```

```
RESULT 3
P90433 SIVCZ
ID P90433_SIVCZ PRELIMINARY; PRT; 313 AA.
AC P90433;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Truncated reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Smith J.M., Kraiselburd E.N., Torres J.V.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U83413; AAB41428.1; -; Genomic_DNA.
DR HSSP; Q07387; 1TCW.
DR SMR; P90433; 69-167, 173-313.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
```



```
DR PROSITE; PS50878; RT POL; 1.
DR PROSITE; PS50876; ZF INTEGRASE; 1.
KW AIDS; Aspartyl protease; DNA integration; DNA recombination;
KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
KW Nuclease; Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferase; Zinc; Zinc-finger.
FT CHAIN 1 167 Protease.
FT DOMAIN 88 157 Peptidase A2.
FT DOMAIN 211 401 Reverse transcriptase.
FT DOMAIN 600 723 RNase H.
FT DOMAIN 779 930 Integrase catalytic.
FT ZN_FING 729 770 Integrase-type.
FT DNA_BIND 949 996 Integrase-type.
FT ACT_SITE 93 93 By similarity.
SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match 75.0%; Score 42; DB 1; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
||| ||||
Db 184 EGPKLRQW 191

RESULT 6
P89154 SIVCZ
ID P89154_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC P89154;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Pol protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SIVsmE543;
RX MEDLINE=97151152; PubMed=8995688;
RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,
RA Elkins W.R., Montefiori D.C.;
RT "A molecularly cloned, pathogenic, neutralization-resistant simian
RT immunodeficiency virus, SIVsmE543-3.";
RL J. Virol. 71:1608-1620(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SIVsmE543;
RA Ourmanov I.K., Deghani H., Kuwata T., Hirsch V.M.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -I- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U72748; AAC56559.2; -; Genomic_DNA.
DR SMR; P89154; 69-167, 781-939.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
```

```
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1 1
SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
||| ||||
Db 184 EGPKLRQW 191

RESULT 7
Q7ZBR7 SIVCZ
ID Q7ZBR7_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR7;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Deghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -I- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -I- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221514; AA067307.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
```

DR SMR; Q7ZBR7; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115341 MW; A886525DFF1BE26F CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.le+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
||| ||||
Db 184 EGPKLRQW 191

RESULT 8
Q7ZBR5 SIVCZ
ID Q7ZBR5_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;

RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPOG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221515; AAC67309.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.le+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```
QY      1 EGPTLRQW 8
      ||| ||||
Db      184 EGPKLRQW 191

RESULT 9
Q9R7K1_9SPHN
ID      Q9R7K1_9SPHN PRELIMINARY;      PRT;      244 AA.
AC      Q9R7K1;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Photosynthetic reaction center L subunit (Fragment).
GN      Name=pufL;
OS      Erythrobacter sp.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC      Sphingomonadaceae; Erythrobacter.
OX      NCBI_TaxID=1042;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=MBIC3031;
RX      MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA      Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA      Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT      "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL      Nature 415:630-633(2002).
DR      EMBL; AB015710; BAA32999.1; -; Genomic_DNA.
DR      HSSP; P02954; 1YST.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR      GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR      GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR      GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR      GO; GO:0000287; F:magnesium ion binding; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR      GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR005871; Photo_L.
DR      InterPro; IPR000484; Photo_RC.
DR      Pfam; PF00124; Photo_RC; 1.
DR      PRINTS; PR00256; REACTNCENTRE.
DR      TIGRFAMs; TIGR01157; pufL; 1.
DR      PROSITE; PS00244; REACTION_CENTER; 1.
FT      NON TER      1
SQ      SEQUENCE      244 AA; 27110 MW; 1B70EC636A39D95A CRC64;

Query Match      73.2%; Score 41; DB 2; Length 244;
Best Local Similarity      77.8%; Pred. No. 39;
Matches      7; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

QY      1 EGPTLRQWL 9
      ||||| ||
Db      22 EGPTLNPWL 30

RESULT 10
O66272_9SPHN
ID      O66272_9SPHN PRELIMINARY;      PRT;      245 AA.
AC      O66272;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Photosynthetic reaction center L subunit (Fragment).
GN      Name=pufL;
OS      Erythrobacter litoralis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC      Sphingomonadaceae; Erythrobacter.
OX      NCBI_TaxID=39960;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=IAM14332;
RX      MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;

Query Match      73.2%; Score 41; DB 2; Length 244;
Best Local Similarity      77.8%; Pred. No. 39;
Matches      7; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

QY      1 EGPTLRQWL 9
      ||||| ||
Db      22 EGPTLNPWL 30

RESULT 11
O82991_9SPHN
ID      O82991_9SPHN PRELIMINARY;      PRT;      245 AA.
AC      O82991;
DT      01-NOV-1998 (TrEMBLrel. 08, Created)
DT      01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Photosynthetic reaction center L subunit (Fragment).
GN      Name=pufL;
OS      Erythrobacter sp.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC      Sphingomonadaceae; Erythrobacter.
OX      NCBI_TaxID=1042;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=MBIC3017;
RX      MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA      Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA      Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT      "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL      Nature 415:630-633(2002).
DR      EMBL; AB015709; BAA32997.1; -; Genomic_DNA.
DR      HSSP; P02954; 1YST.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR      GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR      GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR      GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR      GO; GO:0000287; F:magnesium ion binding; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR      GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR005871; Photo_L.
DR      InterPro; IPR000484; Photo_RC.
DR      Pfam; PF00124; Photo_RC; 1.
DR      PRINTS; PR00256; REACTNCENTRE.
DR      TIGRFAMs; TIGR01157; pufL; 1.
DR      PROSITE; PS00244; REACTION_CENTER; 1.
FT      NON TER      1
SQ      SEQUENCE      244 AA; 27110 MW; 1B70EC636A39D95A CRC64;
```

```
RA      Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA      Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT      "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL      Nature 415:630-633(2002).
DR      EMBL; AB010981; BAA25791.1; -; Genomic_DNA.
DR      HSSP; P02954; 1QOV.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR      GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR      GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR      GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR      GO; GO:0000287; F:magnesium ion binding; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR      GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR005871; Photo_L.
DR      InterPro; IPR000484; Photo_RC.
DR      Pfam; PF00124; Photo_RC; 1.
DR      PRINTS; PR00256; REACTNCENTRE.
DR      TIGRFAMs; TIGR01157; pufL; 1.
DR      PROSITE; PS00244; REACTION_CENTER; 1.
FT      NON TER      1
SQ      SEQUENCE      245 AA; 27214 MW; 52B268713E199ABD CRC64;

Query Match      73.2%; Score 41; DB 2; Length 245;
Best Local Similarity      77.8%; Pred. No. 39;
Matches      7; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

QY      1 EGPTLRQWL 9
      ||||| ||
Db      27 EGPTLNPWL 35

RESULT 11
O82991_9SPHN
ID      O82991_9SPHN PRELIMINARY;      PRT;      245 AA.
AC      O82991;
DT      01-NOV-1998 (TrEMBLrel. 08, Created)
DT      01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Photosynthetic reaction center L subunit (Fragment).
GN      Name=pufL;
OS      Erythrobacter sp.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC      Sphingomonadaceae; Erythrobacter.
OX      NCBI_TaxID=1042;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=MBIC3017;
RX      MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA      Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA      Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT      "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL      Nature 415:630-633(2002).
DR      EMBL; AB015709; BAA32997.1; -; Genomic_DNA.
DR      HSSP; P02954; 1YST.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR      GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR      GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR      GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR      GO; GO:0000287; F:magnesium ion binding; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR      GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR005871; Photo_L.
DR      InterPro; IPR000484; Photo_RC.
DR      Pfam; PF00124; Photo_RC; 1.
DR      PRINTS; PR00256; REACTNCENTRE.
DR      TIGRFAMs; TIGR01157; pufL; 1.
DR      PROSITE; PS00244; REACTION_CENTER; 1.
DR
```


QY 2 GPTLRQWL 9
|||
Db 286 GPDLRQWL 293

RESULT 15
XERC_CORGL STANDARD; PRT; 308 AA.
AC Q8NNZ9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Tyrosine recombinase xerC.
GN Name=xerC; OrderedLocusNames=Cgl2028, cg2224;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puehler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegraebe I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
and its impact on the production of L-aspartate-derived amino acids
and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
CC -|- FUNCTION: Site-specific tyrosine recombinase, which acts by
catalyzing the cutting and rejoining of the recombining DNA
molecules. The xerC-xerD complex is essential to convert dimers of
the bacterial chromosome into monomers to permit their segregation
at cell division. It also contributes to the segregational
stability of plasmids (By similarity).
CC -|- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two
molecules of xerC and two molecules of xerD (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the "phage" integrase family. XerC type 1
subfamily.

This Swiss-prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; BA000036; BAB99421.1; -; Genomic DNA.
DR EMBL; BX927154; CAF20368.1; ALT_INIT; Genomic DNA.
DR HSSP; P21891; 1AOP.
DR HAMAP; MF 01808; -; 1.
DR InterPro; IPR004107; Phage_integr_N.
DR InterPro; IPR002104; Phage_integrase.
DR Pfam; PF02899; Phage_integr_N; 1.
DR Pfam; PF00589; Phage_integrase; 1.
KW Cell cycle; Cell division; Chromosome partition; Complete proteome;
KW DNA integration; DNA recombination; DNA-binding.
FT ACT_SITE 163 163 By similarity.
FT ACT_SITE 187 187 By similarity.
FT ACT_SITE 254 254 By similarity.
FT ACT_SITE 257 257 By similarity.
FT ACT_SITE 280 280 By similarity.
FT ACT_SITE 289 289 O-(3'-phospho-DNA)-tyrosine intermediate
(By similarity).
FT

SQ SEQUENCE 308 AA; 34053 MW; 10PA3E71949334FC CRC64;
Query Match 73.2%; Score 41; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRQWL 9
|||||
Db 67 PTLRQWL 73

RESULT 16
Q617Q3 CAEBR PRELIMINARY; PRT; 728 AA.
AC Q617Q3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14874.
GN Name=CBG14874;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; CAAC01000068; CAE68902.1; -; Genomic DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Ubl conjugation pathway; Zinc;
KW Zinc-finger.
SQ SEQUENCE 728 AA; 82017 MW; 97016C238A26F96B CRC64;
Query Match 73.2%; Score 41; DB 2; Length 728;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
|||
Db 614 GPCLRKWLA 622

RESULT 17
Q95Y82 CAEEL PRELIMINARY; PRT; 754 AA.
AC Q95Y82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Y119C1B.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).


```
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AC006712; AAK39324.1; -; Genomic DNA.
DR Ensembl; Y119C1B.5; Caenorhabditis elegans.
DR WormBase; WBGene00022471; Y119C1B.5.
DR WormPep; Y119C1B.5; CE27234.
DR DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR DR GO; GO:0046872; F:metal ion binding; IEA.
DR DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR DR GO; GO:0008270; F:zinc ion binding; IEA.
DR DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR DR InterPro; IPR001841; Znf_RING.
DR DR Pfam; PF00097; zf-C3HC4; 1.
DR DR SMART; SM00184; RING; 1.
DR DR PROSITE; PS00089; ZF_RING_2; 1.
KW Complete proteome; Hypothetical protein; Metal-binding;
KW Ubl conjugation pathway; Zinc; Zinc-finger.
SQ SEQUENCE 754 AA; 85324 MW; 41BAA9297FA3BF05 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 754;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
Db 626 GPCLRKWLA 634

RESULT 18
Q73SJ6_MYCPA
ID Q73SJ6_MYCPA PRELIMINARY; PRT; 791 AA.
AC Q73SJ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP4077c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017242; AAS06627.1; -; Genomic_DNA.
DR InterPro; IPR005242; Cons_hypoth374.
DR Pfam; PF03706; UPF0104; 1.
DR TIGRFAMs; TIGR00374; Cons_hypoth374; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 791 AA; 83885 MW; 2E5CB0DE4B2B4F1B CRC64;

Query Match 73.2%; Score 41; DB 2; Length 791;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
Db 653 PTLRRWLA 660

RESULT 19
Q4KS46_9VIRU
ID Q4KS46_9VIRU PRELIMINARY; PRT; 815 AA.
AC Q4KS46;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tyrosine kinase.
OS Orange-spotted grouper iridovirus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=322017;
```

```
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhou S.Y., Lv L., Chen C., Weng S.P., Chan S.M., He J.G.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY894343; AAX82420.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 815 AA; 93153 MW; 120C2FB59703F8C4 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 815;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
Db 501 QGPTLAQWI 509

RESULT 20
Q9ST50_MAIZE
ID Q9ST50_MAIZE PRELIMINARY; PRT; 863 AA.
AC Q9ST50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transposase related protein.
GN Name=trap;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Baumeister P., Koenig J., Werr W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ238507; CAB51950.1; -; Genomic_DNA.
DR InterPro; IPR004332; MuDR.
DR InterPro; IPR006564; Znf_PMZ.
DR InterPro; IPR007527; Znf_SWIM.
DR Pfam; PF03108; MuDR; 1.
DR Pfam; PF04434; SWIM; 1.
DR SMART; SM00575; Znf_PMZ; 1.
DR PROSITE; PS00966; ZF_SWIM; 1.
DR SEQUENCE 863 AA; 98714 MW; 2B3FDA5C57B172A1 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRQWL 9
Db 305 PTLRQWL 311

RESULT 21
Q8QUJ6_9VIRU
ID Q8QUJ6_9VIRU PRELIMINARY; PRT; 941 AA.
AC Q8QUJ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF114L.
OS Infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Megalocytivirus.
OX NCBI_TaxID=180170;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus.";
RT Virology 291:126-139(2001).
OX NCBI_TaxID=126-139(2001).
```

DR EMBL; AF371960; AAL98838.1; -; Genomic DNA.
SQ SEQUENCE 941 AA; 106703 MW; EB663998C7F6CE83 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 941;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
Db 582 QGPTLAQWI 590
:|||||
:

RESULT 22
Q5YUJ5_NOCFA
ID Q5YUJ5_NOCFA PRELIMINARY; PRT; 53 AA.
AC Q5YUJ5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=nfa25990;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD57446.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 53 AA; 5806 MW; 47B08B1F9BA4FCE6 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 53;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
Db 33 GPQCROWLA 41
|||

RESULT 23
Q8DHX7_SYNEL
ID Q8DHX7_SYNEL PRELIMINARY; PRT; 129 AA.
AC Q8DHX7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tll1816 protein.
GN OrderedLocusNames=tll1816;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Ikeguchi T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Irinouchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; BA000039; BAC09368.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 129 AA; 14644 MW; EBB44691E7DD1E12 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 129;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
Db 42 GATLRQWLS 50
|

RESULT 24
Q6FE11_ACIAD
ID Q6FE11_ACIAD PRELIMINARY; PRT; 137 AA.
AC Q6FE11;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=ACIAD0794;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
DR EMBL; CR543861; CAG67697.1; -; Genomic_DNA.
DR InterPro; IPR007401; DUF454.
DR Pfam; PF04304; DUF454; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 137 AA; 15749 MW; 70C5D7C68908AFCA CRC64;

Query Match 71.4%; Score 40; DB 2; Length 137;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
Db 57 GPTLRDWRA 65
|

RESULT 25
Q54XF9_DICDI
ID Q54XF9_DICDI PRELIMINARY; PRT; 154 AA.
AC Q54XF9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DDB0215283;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Tunggul B., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,

RA Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFI01000054; EAL67916.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 154 AA; 17998 MW; 2BE3834C15F6FAF6 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 154;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
Db 73 QGPTSKQWL 81

RESULT 26
Q8ZGS7 YERPE
ID Q8ZGS7_YERPE PRELIMINARY; PRT; 296 AA.
AC Q8ZGS7; Q74WE0; Q7CH89;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative membrane protein (Putative transmembrane protein).
GN Name=rha74; OrderedLocusNames=YP0934, YPO1203, Y2985;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V., Blattner F.R.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";

RL DNA Res. 11:179-197(2004).
DR EMBL; AJ414147; CAC90042.1; -; Genomic_DNA.
DR EMBL; AE013900; AAM86536.1; -; Genomic_DNA.
DR EMBL; AE017130; AAS61189.1; -; Genomic_DNA.
DR PIR; AG0147; AG0147.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 296 AA; 31378 MW; 45947413DCD54CFF CRC64;

Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
Db 66 PTLRQWAA 73

RESULT 27
Q66D06 YERPS
ID Q66D06_YERPS PRELIMINARY; PRT; 296 AA.
AC Q66D06;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative drug/metabolite (DME family) efflux pump precursor.
GN OrderedLocusNames=YPTB1243;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH20483.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 25 Potential.
SQ SEQUENCE 296 AA; 31407 MW; 4D3E486D32DBAC11 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
Db 66 PTLRQWAA 73

RESULT 28
P95613 RHIGA
ID P95613_RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAMBI;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstroem K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.
DR EMBL; Y08963; CAA70157.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH 1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS0931; HTH_LysR; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 71.4%; Score 40; DB 2; Length 326;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
Db 204 KGPSLEQWLS 213
:::|::|::|

RESULT 29
Q82PX5 STRAW
ID Q82PX5_STRAW PRELIMINARY; PRT; 377 AA.
AC Q82PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV747;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.2114331198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC68457.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 377 AA; 41307 MW; 0253176AAAB562F3 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 377;
Best Local Similarity 77.8%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
Db 169 EGPDRAWL 177
::|::|::|

RESULT 30
Q7S2Z7 NEUCR
ID Q7S2Z7_NEUCR PRELIMINARY; PRT; 429 AA.
AC Q7S2Z7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU07552.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Biele C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000415; EAA29829.1; -; Genomic_DNA.
DR GO; GO:0005488; F:binding; IEA.
DR InterPro; IPR011990; TPR-like_helical.
SQ SEQUENCE 429 AA; 50079 MW; 5C85703F66E71B8 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 429;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
Db 16 QGPTLEQW 23
:|::|::|

RESULT 31
Q7G4N4 ORYSA
ID Q7G4N4_ORYSA PRELIMINARY; PRT; 760 AA.
AC Q7G4N4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Putative gag-pol polyprotein (Hypothetical protein).
GN ORFNames=OSJNBb0038H12.4;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";

```

RL Science 300:1566-1569(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,
RA Bera J., Kim M., Jin S., Fadrosch D., Vuong H., Overton II L.,
RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,
RA de Vazeilles A., White O., Salzberg S., Fraser C.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017067; AAP52546.1; -; Genomic_DNA.
DR EMBL; AC105932; AAX95563.1; -; Genomic_DNA.
DR Gramene; Q8LMK9; -;
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein; Polyprotein.
SQ SEQUENCE 760 AA; 82020 MW; C51F91AA2EB32A28 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 760;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
   ||||:|:|
Db 663 GPTLQHWMA 671

RESULT 32
P72684_SVNY3
ID P72684_SVNY3 PRELIMINARY; PRT; 131 AA.
AC P72684;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Slr0740 protein.
GN OrderedLocusNames=slr0740;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; BA000022; BAA16691.1; -; Genomic_DNA.
DR PIR; S74539; S74539.
KW Complete proteome.
SQ SEQUENCE 131 AA; 14495 MW; DB0B8997ACF2CB70 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRQWLA 10
   |||||
Db 29 PQLRQWLA 36
```

```

RESULT 33
Q4ZPV7_PSESY
ID Q4ZPV7_PSESY PRELIMINARY; PRT; 133 AA.
AC Q4ZPV7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Psyf_3784;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
RA Nolan M., Goltsman E., Thiel J., Malfatti S., Lapidus A., Detter J.C.,
RA Land M., Richardson P.M., Kyrpides N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
RT pv. syringae B728a and pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Loper J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Feil H., Feil W.S., Lindow S.E.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000075; AAY38815.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 133 AA; 14404 MW; 4176926884998F43 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 133;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
   ||||:|:|
Db 28 GPTYREWLV 35

RESULT 34
Q88615_PSESM
ID Q88615_PSESM PRELIMINARY; PRT; 134 AA.
AC Q88615;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PSPT01594;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidsen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
```

```
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016853; AAO55114.1; -; Genomic_DNA.
DR TIGR; PSPTO1594; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 134 AA; 14322 MW; 81A064617211C28E CRC64;

Query Match 69.6%; Score 39; DB 2; Length 134;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
   |||||
Db 28 GPTYREWL 35

RESULT 35
Q9WI01_9HIV2
ID Q9WI01_9HIV2 PRELIMINARY; PRT; 154 AA.
AC Q9WI01;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protease (Fragment).
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20211658; PubMed=10747109;
RA Rodas B., Holguin A., Soriano V., Dourana M., Mansinho K., Antunes F.,
RA Gonzalez-Lahoz J.;
RT "Emergence of drug resistance mutations in human immunodeficiency
RT virus type 2-infected subjects undergoing antiretroviral therapy.";
RL J. Clin. Microbiol. 38:1370-1374(2000).
DR EMBL; AF139053; AAD37335.1; -; Genomic_DNA.
DR HSSP; P04584; LIDA.
DR SMR; Q9WI01; 1-98, 101-154.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 154
FT NON_TER 154
SQ SEQUENCE 154 AA; 16821 MW; C47EB4C660405791 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 154;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
   :|||
Db 115 DGPKLRQW 122

RESULT 36
Q6MNC9_BDEBA
ID Q6MNC9_BDEBA PRELIMINARY; PRT; 164 AA.
AC Q6MNC9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedLocusNames=Bdi329;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
```

```
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842649; CAE79223.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1
FT SIGNAL 7 Potential.
SQ SEQUENCE 164 AA; 17624 MW; D9B7CCFDDA0246A2 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 164;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWLA 10
   :|||
Db 60 QGPALQWTA 69

RESULT 37
Q9V492_DROME
ID Q9V492_DROME PRELIMINARY; PRT; 168 AA.
AC Q9V492;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG11077-PA (RE55125p).
GN ORFNames=CG11077;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
```


RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celniker S.E.;
RA "The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective";
RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R., Hoskins R., Stapleton M., Pacleeb J., Park S., Svirskas R., Smith E., Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G., Miranda A., Mungall C.J., Nunoo J., Pacleeb J., Paragas V., Park S., Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003846; AAF59393.1; -; Genomic_DNA.
DR EMBL; AY071482; AAL49104.1; -; mRNA.
DR Ensembl; CG11077; Drosophila melanogaster.
DR FlyBase; FBgn0039930; CG11077.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR007248; Mpv17_PMP22.
DR Pfam; PF04117; Mpv17_PMP22; 1.
SQ SEQUENCE 168 AA; 19521 MW; 48E216A954E43D39 CRC64;

Query Match

69.6%; Score 39; DB 2; Length 168;

Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 GPTLRQW 8
Db |||||:
54 GPTLRRW 60
RESULT 38
Q87115_SIVCZ
ID Q87115_SIVCZ PRELIMINARY; PRT; 217 AA.
AC Q87115;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SIVagmsABD37;
RX MEDLINE=94298785; PubMed=8026477;
RA Jin M.J., Hui H., Robertson D.L., Muller M.C., Barre-Sinoussi F., Hirsch V.M., Allan J.S., Shaw G.M., Sharp P.M., Hahn B.H.;
RT "Mosaic genome structure of simian immunodeficiency virus from west African green monkeys";
RL EMBO J. 13:2935-2947(1994).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U04018; AAA21512.1; -; Genomic_DNA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR009007; Pept Aspartc cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50878; RT_FOL; 1.
FT NON_TER 1 217
FT NON_TER 217 217
SQ SEQUENCE 217 AA; 24503 MW; C1162E4BE18204B8 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
Db :||| ||||
86 DGPRLRQW 93

RESULT 39
Q5RJN8_RAT
ID Q5RJN8_RAT PRELIMINARY; PRT; 218 AA.
AC Q5RJN8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical LOC287938.
GN Name=LOC287938;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC086565; AAH86565.1; -; mRNA.
DR InterPro; IPR006879; YdJC.
DR Pfam; PF04794; YdJC; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 23436 MW; 7BCCAF3DD2B6ABCE CRC64;

Query Match 69.6%; Score 39; DB 2; Length 218;
Best Local Similarity 87.5%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLQWLA 10
Db 188 PTLRAWLA 195
|||||

RESULT 40
Q7ZKL1_9HIV1 PRELIMINARY; PRT; 230 AA.
AC Q7ZKL1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 1.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22866097; PubMed=14506786; DOI=10.1089/0892220332280946;
RA Toni T.D., Recordon-Pinson P., Minga A., Ekouevi D., Bonard D.,
RA Bequet L., Huet C., Chenal H., Rouet F., Dabis F., Lafon M.-E.,
RA Salamon R., Masquelier B., Fleury H.J.;
RT "Presence of key drug resistance mutations in isolates from untreated
patients of Abidjan, Cote d'Ivoire : ANRS 1257 Study.";
RL AIDS Res. Hum. Retroviruses 19:713-717(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fleury H., Toni T., Recordon-Pinson P., Minga A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY207863; AAO53469.1; -; Genomic_DNA.
DR SNR; Q7ZKL1; 1-230.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 230 230
SQ SEQUENCE 230 AA; 26696 MW; 0FBDBA8B46D5952C CRC64;

Query Match 69.6%; Score 39; DB 2; Length 230;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
Db 3 DGPTVKQW 10
:|||:|

RESULT 41
O82987_9SPHN PRELIMINARY; PRT; 245 AA.
AC O82987;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3361;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB015707; BAA32993.1; -; Genomic_DNA.
DR HSSP; P02954; 1PSS.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 245 AA; 27411 MW; 06C24DD9610CF92D CRC64;

Query Match 69.6%; Score 39; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
Db 22 EGPTMNPWL 30
|||||

RESULT 42
Q5YPT6_NOCFA PRELIMINARY; PRT; 298 AA.
ID Q5YPT6_NOCFA
AC Q5YPT6;

DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Putative UTP-glucose-1-phosphate uridylyltransferase.
GN Name=galu; OrderedLocusNames=nfa49530;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RA "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
RL EMBL; AP006618; BAD59805.1; -; Genomic_DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 298 AA; 31931 MW; F29B1A16B26ABA0D CRC64;

Query Match 69.6%; Score 39; DB 2; Length 298;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
Db 275 GPALREWL 282
[1] |||||
[2] |||||

RESULT 43
Q9D8Z7 MOUSE
ID Q9D8Z7_MOUSE PRELIMINARY; PRT; 310 AA.
AC Q9D8Z7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched
DE library, clone:1810015A11 product:ubiquitin-conjugating enzyme E2L 3,
DE full insert sequence.
DE Name=1810015A11Rik;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipillar sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007503; BAB25075.1; -; mRNA.
DR Ensembl; ENSMUSG00000041774; Mus musculus.
DR MGI; MGI:1916351; 1810015A11Rik.
DR InterPro; IPR006879; YdjC.
DR Pfam; PF04794; YdjC; 1.
SQ SEQUENCE 310 AA; 33094 MW; 9D3302745CC24048 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 310;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLROWLA 10
Db 280 PTLRAWLA 287
[1] |||||
[2] |||||

RESULT 44

Q7SKI9 9HIV2
ID Q7SKI9_9HIV2 PRELIMINARY; PRT; 310 AA.
AC Q7SKI9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/088922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveusson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
RT antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324219; AAP85282.1; -; Genomic_DNA.
DR HSSP; P04584; 1IDA.
DR SMR; Q7SKI9; 1-99, 105-310.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT.1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_FOL; 1.
KW Aspartyl protease; Hydrolase; Polyprotein; Protease.
FT NON_TER 1
FT NON_TER 310 310
SQ SEQUENCE 310 AA; 35077 MW; 33816141D95E1FD1 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 310;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
:|||||
Db 116 DGPKLRQW 123

RESULT 45
Q4PIZ3 USTMA
ID Q4PIZ3_USTMA PRELIMINARY; PRT; 338 AA.
AC Q4PIZ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05870.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,

Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Doris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Phunkhang P., Piquani B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sougnez C.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000215; EAK86815.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 37245 MW; 838DC845D4ACC448 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 338;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
||:|:|
Db 294 GPSLRKWL 301

RESULT 46
Q8UN04 SIVCZ
ID Q8UN04_SIVCZ PRELIMINARY; PRT; 340 AA.
AC Q8UN04;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF458219; AAL59619.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q8UN04; 3-340.

DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 340 340
SQ SEQUENCE 340 AA; 39546 MW; F9F3BFD3F4005252 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
: || || || ||
Db 17 DGPKLQW 24

RESULT 47
Q8UN03_SIVCZ
ID Q8UN03_SIVCZ PRELIMINARY; PRT; 340 AA.
AC Q8UN03;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF458220; AAL59620.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q8UN03; 3-340.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 340 340
SQ SEQUENCE 340 AA; 39548 MW; 7777BFD2A057EA6B CRC64;

Query Match 69.6%; Score 39; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
: || || || ||
Db 17 DGPKLQW 24

RESULT 48
Q7SKK9_9HIV2
ID Q7SKK9_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).

GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/088922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveesson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324199; AAP85262.1; -; Genomic_DNA.
DR HSSP; P04584; 1IDA.
DR SMR; Q7SKK9; 1-99, 102-349.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Polyprotein; Protease.
FT NON_TER 1 1
FT NON_TER 349 349
SQ SEQUENCE 349 AA; 39974 MW; 2B66AE503EFC35B8 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
: || || || ||
Db 116 DGPKLQW 123

RESULT 49
Q7SKK8_9HIV2
ID Q7SKK8_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/088922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveesson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324200; AAP85263.1; -; Genomic_DNA.
DR HSSP; P04584; 1IDA.
DR SMR; Q7SKK8; 1-99, 102-349.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_FOL; 1.
KW Aspartyl protease; Hydrolase; Polyprotein; Protease.
FT NON_TER 1
FT NON_TER 349
SQ SEQUENCE 349 AA; 39855 MW; 82DAEF2DC13E6DED CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
Db 116 DGPKLQW 123

RESULT 50
Q7SKK6_9HIV2
ID Q7SKK6_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/088922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveesson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324202; AAP85265.1; -; Genomic_DNA.
DR HSSP; P04584; IIDA.
DR SMR; Q7SKK6; 1-99, 102-349.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_FOL; 1.
KW Aspartyl protease; Hydrolase; Polyprotein; Protease.
FT NON_TER 1
FT NON_TER 349
SQ SEQUENCE 349 AA; 39830 MW; 593B83A7F9FFA8D0 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;

Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
Db 116 DGPKLQW 123

Search completed: May 12, 2006, 10:51:45
Job time : 87.7692 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 101.709 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-67
Perfect score: 81
Sequence: 1 CELVGPSLMSWLTG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	14	2	Aaw09483 Thrombopo
2	81	100.0	14	2	Aaw36634 Thrombopo
3	81	100.0	14	2	Aaw36767 Thrombopo
4	81	100.0	14	3	Aab17011 TPO-mimet
5	81	100.0	14	4	Aau25853 Human thr
6	81	100.0	14	5	Abb72897 TPO mimet
7	81	100.0	14	7	Adj73048 TPO mimet
8	81	100.0	14	8	Adj52683 CH1 delet
9	81	100.0	14	8	Adj51644 CH1 delet
10	53	65.4	18	7	Adn59675 Thrombopo
11	53	65.4	25	7	Adn59730 Thrombopo
12	52	64.2	14	3	Aab17010 TPO-mimet
13	52	64.2	14	5	Abb72896 TPO mimet
14	52	64.2	14	7	Adj73047 TPO mimet
15	52	64.2	14	8	Adj52682 CH1 delet
16	52	64.2	14	8	Adj51643 CH1 delet
17	49	60.5	129	5	Abb09517 Human RNA
18	49	60.5	198	4	Aam29099 Peptide #
19	49	60.5	199	4	Aau27672 Human pro
20	49	60.5	199	4	Aag77853 Human rib
21	49	60.5	199	4	Aau29085 Human PRO
22	49	60.5	199	4	Aag89149 Human sec
23	49	60.5	199	5	Abp43532 Human sec
24	49	60.5	199	5	Aau12076 Human rib

25	49	60.5	199	6	ABU58693	Abu58693 Human PRO
26	49	60.5	199	6	ABU88241	Abu88241 Novel hum
27	49	60.5	199	6	ABU84556	Abu84556 Human sec
28	49	60.5	199	6	ABR66430	Abr66430 Human sec
29	49	60.5	199	6	ABR65820	Abr65820 Human sec
30	49	60.5	199	6	ABU99760	Abu99760 Human sec
31	49	60.5	199	6	ABU82999	Abu82999 Human PRO
32	49	60.5	199	6	ABU90120	Abu90120 Novel hum
33	49	60.5	199	6	ABR68369	Abr68369 Human sec
34	49	60.5	199	6	ABU96422	Abu96422 Novel hum
35	49	60.5	199	6	ABU92853	Abu92853 Human sec
36	49	60.5	199	6	ABO08930	Abo08930 Human sec
37	49	60.5	199	6	ABO02982	Abo02982 Human sec
38	49	60.5	199	6	ABR75136	Abr75136 Human sec
39	49	60.5	199	6	ABR94898	Abr94898 Human sec
40	49	60.5	199	6	ABU85871	Abu85871 Human PRO
41	49	60.5	199	6	ABU99031	Abu99031 Novel hum
42	49	60.5	199	6	ABU98246	Abu98246 Novel hum
43	49	60.5	199	6	ABU91952	Abu91952 Novel hum
44	49	60.5	199	6	ABU89645	Abu89645 Human PRO
45	49	60.5	199	6	ABU86486	Abu86486 Human sec
46	49	60.5	199	6	ABU67699	Abu67699 Human sec
47	49	60.5	199	6	ABU80727	Abu80727 Human PRO
48	49	60.5	199	6	ABR99645	Abr99645 Human sec
49	49	60.5	199	6	ABR99035	Abr99035 Human sec
50	49	60.5	199	6	ABO16558	Abo16558 Human sec
51	49	60.5	199	6	ABR92458	Abr92458 Human sec
52	49	60.5	199	6	ABO19099	Abo19099 Human sec
53	49	60.5	199	6	ABR78520	Abr78520 Human sec
54	49	60.5	199	6	ABU85256	Abu85256 Novel hum
55	49	60.5	199	6	ABO00395	Abo00395 Novel hum
56	49	60.5	199	6	ABO11727	Abo11727 Human sec
57	49	60.5	199	6	ABO02372	Abo02372 Human sec
58	49	60.5	199	6	ABU88946	Abu88946 Novel hum
59	49	60.5	199	6	ABU83641	Abu83641 Human sec
60	49	60.5	199	6	ABO06442	Abo06442 Novel hum
61	49	60.5	199	6	ABR59478	Abr59478 Human sec
62	49	60.5	199	6	ABO09540	Abo09540 Human sec
63	49	60.5	199	6	ABO19404	Abo19404 Novel hum
64	49	60.5	199	6	ABO11422	Abo11422 Human sec
65	49	60.5	199	6	ABR67040	Abr67040 Human sec
66	49	60.5	199	6	ABO16253	Abo16253 Human sec
67	49	60.5	199	6	ABO13959	Abo13959 Human sec
68	49	60.5	199	6	ABU65862	Abu65862 Human sec
69	49	60.5	199	6	ABO07710	Abo07710 Human PRO
70	49	60.5	199	6	ABO03897	Abo03897 Human sec
71	49	60.5	199	6	ABR67345	Abr67345 Human sec
72	49	60.5	199	6	ABO15948	Abo15948 Human sec
73	49	60.5	199	6	ABU56229	Abu56229 Human sec
74	49	60.5	199	6	ABU65557	Abu65557 Human PRO
75	49	60.5	199	6	ABU95502	Abu95502 Novel hum
76	49	60.5	199	6	ABU71405	Abu71405 Human PRO
77	49	60.5	199	6	ABO08015	Abo08015 Human PRO
78	49	60.5	199	6	ABR70256	Abr70256 Human sec
79	49	60.5	199	6	ABR69589	Abr69589 Human sec
80	49	60.5	199	6	ABO01730	Abo01730 Human PRO
81	49	60.5	199	6	ABU81532	Abu81532 Human PRO
82	49	60.5	199	6	ABR60329	Abr60329 Human sec
83	49	60.5	199	6	ABR68064	Abr68064 Human sec
84	49	60.5	199	6	ABR65452	Abr65452 Human sec
85	49	60.5	199	6	ABR68674	Abr68674 Human sec
86	49	60.5	199	6	ABR72086	Abr72086 Human sec
87	49	60.5	199	6	ABU85566	Abu85566 Human PRO
88	49	60.5	199	6	ABU89256	Abu89256 Human sec
89	49	60.5	199	6	ABU83336	Abu83336 Human sec
90	49	60.5	199	6	ABU95192	Abu95192 Novel hum
91	49	60.5	199	6	ABU90740	Abu90740 Novel hum
92	49	60.5	199	6	ABU84251	Abu84251 Human sec
93	49	60.5	199	6	ABU93902	Abu93902 Novel hum
94	49	60.5	199	6	ABR65147	Abr65147 Human sec
95	49	60.5	199	6	ABR68979	Abr68979 Human sec
96	49	60.5	199	6	ABO06795	Abo06795 Human sec
97	49	60.5	199	6	ABR99340	Abr99340 Human sec

98	49	60.5	199	6	ABU57224	Human	PRO	171	49	60.5	199	6	ABM35307	Human	sec	Abm35307
99	49	60.5	199	6	ABU86176	Novel	hum	172	49	60.5	199	6	ABM26544	Human	sec	Abm26544
100	49	60.5	199	6	ABU82463	Novel	hum	173	49	60.5	199	6	ABO48326	Human	sec	AbO48326
101	49	60.5	199	6	ABU87474	Human	PRO	174	49	60.5	199	6	ABR93068	Human	sec	Abr93068
102	49	60.5	199	6	ABU83946	Human	sec	175	49	60.5	199	6	ABO24829	Human	sec	Abo24829
103	49	60.5	199	6	ABO08320	Human	PRO	176	49	60.5	199	6	ABM11840	Human	sec	Abml1840
104	49	60.5	199	6	ABU82031	Novel	hum	177	49	60.5	199	6	ABM02941	Human	sec	Abm02941
105	49	60.5	199	6	ABU66195	Novel	hum	178	49	60.5	199	6	ABM16237	Human	sec	Abm16237
106	49	60.5	199	6	ABR60024	Human	sec	179	49	60.5	199	6	ABO27798	Human	sec	Abo27798
107	49	60.5	199	6	ABU94212	Novel	hum	180	49	60.5	199	6	ABM29289	Human	sec	Abm29289
108	49	60.5	199	6	ABO00085	Novel	hum	181	49	60.5	199	6	ABM07265	Human	sec	Abm07265
109	49	60.5	199	6	ABR66735	Human	sec	182	49	60.5	199	6	ABM21359	Human	sec	Abm21359
110	49	60.5	199	6	ABR91153	Human	sec	183	49	60.5	199	6	ABM09705	Human	sec	Abm09705
111	49	60.5	199	6	ABU94580	Human	PRO	184	49	60.5	199	6	ABO41575	Human	sec	Abo41575
112	49	60.5	199	6	ABU79462	Human	PRO	185	49	60.5	199	6	ABO36390	Human	PRO	Abo36390
113	49	60.5	199	6	ABU86791	Human	sec	186	49	60.5	199	6	ABO43919	Human	PRO	Abo43919
114	49	60.5	199	6	ABU87096	Novel	hum	187	49	60.5	199	6	ABM76619	Human	sec	Abm76619
115	49	60.5	199	6	ABU94885	Human	PRO	188	49	60.5	199	6	ABM76315	Human	sec	Abm76315
116	49	60.5	199	6	ABO04812	Human	PRO	189	49	60.5	199	6	ABM25934	Human	sec	Abm25934
117	49	60.5	199	6	ABR70561	Human	sec	190	49	60.5	199	6	ABM26239	Human	sec	Abm26239
118	49	60.5	199	6	ABU98726	Human	PRO	191	49	60.5	199	6	ABO03592	Human	sec	Abo03592
119	49	60.5	199	6	ABR66125	Human	sec	192	49	60.5	199	6	ABO02677	Human	sec	Abo02677
120	49	60.5	199	6	ABR64842	Human	sec	193	49	60.5	199	6	ABR90848	Human	sec	Abr90848
121	49	60.5	199	6	ABU79767	Human	PRO	194	49	60.5	199	6	ABR73916	Human	sec	Abr73916
122	49	60.5	199	6	ABU93158	Human	sec	195	49	60.5	199	6	ABO17168	Human	sec	Abo17168
123	49	60.5	199	6	ABU96117	Human	PRO	196	49	60.5	199	6	ABR94593	Human	sec	Abr94593
124	49	60.5	199	6	ABU91337	Novel	hum	197	49	60.5	199	6	ABR76100	Human	sec	Abr76100
125	49	60.5	199	6	ABU90430	Human	hum	198	49	60.5	199	6	ABR71476	Human	sec	Abr71476
126	49	60.5	199	6	ABO09845	Human	sec	199	49	60.5	199	6	ABR93373	Human	sec	Abr93373
127	49	60.5	199	6	ABO11117	Human	sec	200	49	60.5	199	6	ABR93678	Human	sec	Abr93678
128	49	60.5	199	6	ABR71171	Human	sec	201	49	60.5	199	6	ABR88103	Human	sec	Abr88103
129	49	60.5	199	6	ABU87779	Human	PRO	202	49	60.5	199	6	ABO28103	Human	sec	Abo28103
130	49	60.5	199	6	ABU91647	Human	PRO	203	49	60.5	199	6	ABO30238	Human	sec	Abo30238
131	49	60.5	199	6	ABU84861	Human	sec	204	49	60.5	199	6	ABO33447	Human	PRO	Abo33447
132	49	60.5	199	6	ABR69951	Human	sec	205	49	60.5	199	6	ABM05135	Human	sec	Abm05135
133	49	60.5	199	6	ABU80328	Human	PRO	206	49	60.5	199	6	ABM09095	Human	sec	Abm09095
134	49	60.5	199	6	ABU93597	Human	PRO	207	49	60.5	199	6	ABO36695	Human	sec	Abo36695
135	49	60.5	199	6	ABO10150	Human	sec	208	49	60.5	199	6	ABO35780	Human	PRO	Abo35780
136	49	60.5	199	6	ABO09235	Human	sec	209	49	60.5	199	6	ABO39745	Human	sec	Abo39745
137	49	60.5	199	6	ABU10803	Human	sec	210	49	60.5	199	6	ABM10620	Human	sec	Abm10620
138	49	60.5	199	6	ABU95812	Human	PRO	211	49	60.5	199	6	ABM12145	Human	sec	Abm12145
139	49	60.5	199	6	ABU97021	Novel	hum	212	49	60.5	199	6	ABO52291	Human	PRO	Abo52291
140	49	60.5	199	6	ABR70866	Human	sec	213	49	60.5	199	6	ABO52596	Human	PRO	Abo52596
141	49	60.5	199	6	ABO05217	Novel	hum	214	49	60.5	199	6	ABO23914	Human	sec	Abo23914
142	49	60.5	199	6	ABO08625	Human	sec	215	49	60.5	199	6	ABR97400	Human	sec	Abr97400
143	49	60.5	199	6	ABO05832	Human	sec	216	49	60.5	199	6	ABU63384	Human	rib	Abu63384
144	49	60.5	199	6	ABR74221	Human	sec	217	49	60.5	199	6	ABR87188	Human	sec	Abr87188
145	49	60.5	199	6	ABR95813	Human	sec	218	49	60.5	199	6	ABM11230	Human	sec	Abm11230
146	49	60.5	199	6	ABR81110	Human	sec	219	49	60.5	199	6	ABM28374	Human	sec	Abm28374
147	49	60.5	199	6	ABR81415	Human	sec	220	49	60.5	199	6	ABO32373	Human	sec	Abo32373
148	49	60.5	199	6	ABM01111	Human	sec	221	49	60.5	199	6	ABM15500	Human	sec	Abm15500
149	49	60.5	199	6	ABR88713	Human	sec	222	49	60.5	199	6	ABM06655	Human	sec	Abm06655
150	49	60.5	199	6	ABM77534	Human	sec	223	49	60.5	199	6	ABM04466	Human	sec	Abm04466
151	49	60.5	199	6	ABO29018	Human	sec	224	49	60.5	199	6	ABM22579	Human	sec	Abm22579
152	49	60.5	199	6	ABO31763	Human	sec	225	49	60.5	199	6	ABM07875	Human	sec	Abm07875
153	49	60.5	199	6	ABM08180	Human	sec	226	49	60.5	199	6	ABO40965	Human	sec	Abo40965
154	49	60.5	199	6	ABO40660	Human	sec	227	49	60.5	199	6	ABM35612	Human	sec	Abm35612
155	49	60.5	199	6	ABO36085	Human	PRO	228	49	60.5	199	6	ABM33375	Human	sec	Abm33375
156	49	60.5	199	6	ABO44224	Human	PRO	229	49	60.5	199	6	ABO52901	Human	PRO	Abo52901
157	49	60.5	199	6	ADA78340	Human	sec	230	49	60.5	199	6	ABO50461	Human	sec	Abo50461
158	49	60.5	199	6	ABM25019	Human	sec	231	49	60.5	199	6	ABU99455	Human	sec	Abu99455
159	49	60.5	199	6	ABO03287	Human	sec	232	49	60.5	199	6	ABO04507	Human	sec	Abo04507
160	49	60.5	199	6	ABR90543	Human	sec	233	49	60.5	199	6	ABM18677	Human	sec	Abm18677
161	49	60.5	199	6	ABM17457	Human	sec	234	49	60.5	199	6	ABR97705	Human	sec	Abr97705
162	49	60.5	199	6	ABR95203	Human	sec	235	49	60.5	199	6	ABR80805	Human	sec	Abr80805
163	49	60.5	199	6	ABR95508	Human	sec	236	49	60.5	199	6	ABM01416	Human	sec	Abm01416
164	49	60.5	199	6	ABO21746	Human	sec	237	49	60.5	199	6	ABR89018	Human	sec	Abr89018
165	49	60.5	199	6	ABR98010	Human	sec	238	49	60.5	199	6	ABM13670	Human	sec	Abm13670
166	49	60.5	199	6	ABR87798	Human	sec	239	49	60.5	199	6	ABM21054	Human	sec	Abm21054
167	49	60.5	199	6	ABM77839	Human	sec	240	49	60.5	199	6	ABO42185	Human	sec	Abo42185
168	49	60.5	199	6	ABM28069	Human	sec	241	49	60.5	199	6	ABO42795	Human	sec	Abo42795
169	49	60.5	199	6	ABM06350	Human	sec	242	49	60.5	199	6	ABM10315	Human	sec	Abm10315
170	49	60.5	199	6	ABM03856	Human	sec	243	49	60.5	199	6	ABO38830	Human	sec	Abo38830

390	49	60.5	199	7	ABM13060	Abm13060 Human sec	463	49	60.5	199	8	ADJ55226	Adj55226 Human PRO
391	49	60.5	199	7	ABM30814	Abm30814 Human sec	464	49	60.5	199	8	ADJ64997	Adj64997 Human PRO
392	49	60.5	199	7	ABM24714	Abm24714 Human sec	465	49	60.5	199	8	ADM31893	Adm31893 Novel hum
393	49	60.5	199	7	ABO29628	AbO29628 Human sec	466	49	60.5	199	8	ADM36940	Adm36940 Novel hum
394	49	60.5	199	7	ABO31458	AbO31458 Human sec	467	49	60.5	199	8	ADM40745	Adm40745 Novel hum
395	49	60.5	199	7	ABM14585	Abm14585 Human sec	468	49	60.5	199	8	ADN38353	Adn38353 Novel hum
396	49	60.5	199	7	ABM10010	Abm10010 Human sec	469	49	60.5	500	4	ABB62891	Abb62891 Drosophil
397	49	60.5	199	7	ABO39135	AbO39135 Human sec	470	46	56.8	14	3	AAB17016	Aab17016 TPO-mimet
398	49	60.5	199	7	ABM34900	Abm34900 Human sec	471	46	56.8	14	5	ABB72902	Abb72902 TPO mimet
399	49	60.5	199	7	ABO51376	AbO51376 Human sec	472	46	56.8	14	8	ADJ52688	Adj52688 CH1 delet
400	49	60.5	199	7	ABO04202	AbO04202 Human sec	473	46	56.8	14	8	ADJ51649	Adj51649 CH1 delet
401	49	60.5	199	7	ABO10672	AbO10672 Human PRO	474	46	56.8	401	7	ABO67686	Abo67686 Klebsiell
402	49	60.5	199	7	ABR77915	AbR77915 Human sec	475	45	55.6	14	2	AAW09484	Aaw09484 Thrombopo
403	49	60.5	199	7	ABR79125	AbR79125 Human sec	476	45	55.6	14	2	AAW36766	Aaw36766 Thrombopo
404	49	60.5	199	7	ABO24219	AbO24219 Human sec	477	45	55.6	14	2	AAW36635	Aaw36635 Thrombopo
405	49	60.5	199	7	ABR93983	AbR93983 Human sec	478	45	55.6	14	3	AAB17009	Aab17009 TPO-mimet
406	49	60.5	199	7	ABM02026	Abm02026 Human sec	479	45	55.6	14	4	AAU25854	Aau25854 Human thr
407	49	60.5	199	7	ABM78449	Abm78449 Human sec	480	45	55.6	14	5	ABB72895	Abb72895 TPO mimet
408	49	60.5	199	7	ABR90238	AbR90238 Human sec	481	45	55.6	14	7	ADJ73046	Adj73046 TPO mimet
409	49	60.5	199	7	ABM27764	Abm27764 Human sec	482	45	55.6	14	8	ADJ52681	Adj52681 CH1 delet
410	49	60.5	199	7	ABM13365	Abm13365 Human sec	483	45	55.6	14	8	ADJ51642	Adj51642 CH1 delet
411	49	60.5	199	7	ABO32068	AbO32068 Human sec	484	45	55.6	327	2	AAW73380	Aaw73380 CSB2 prot
412	49	60.5	199	7	ABM14280	Abm14280 Human sec	485	45	55.6	384	3	AAB37414	Aab37414 Human sec
413	49	60.5	199	7	ABM08485	Abm08485 Human sec	486	45	55.6	412	4	AAB95879	Aab95879 Human pro
414	49	60.5	199	7	ABO40355	AbO40355 Human sec	487	45	55.6	526	2	AAW73378	Aaw73378 CSB2 prot
415	49	60.5	199	7	ABM74790	Abm74790 Human sec	488	45	55.6	526	4	AAM39051	Aam39051 Human pol
416	49	60.5	199	7	ABR85358	AbR85358 Human sec	494	44	54.3	18	7	ADN59673	Adn59673 Thrombopo
422	49	60.5	199	7	ABO15338	AbO15338 Human sec	495	44	54.3	22	7	ADN59840	Adn59840 TMP pepti
423	49	60.5	199	7	ABO17473	AbO17473 Human sec	496	44	54.3	22	7	ADN59841	Adn59841 TMP pepti
424	49	60.5	199	7	ABM17762	Abm17762 Human sec	497	44	54.3	23	7	ADN59797	Adn59797 Peptide-v
425	49	60.5	199	7	ABR85663	AbR85663 Human sec	498	44	54.3	23	7	ADN59778	Adn59778 Peptide-v
426	49	60.5	199	7	ABM77229	Abm77229 Human sec	499	44	54.3	23	7	ADN59779	Adn59779 Peptide-v
427	49	60.5	199	7	ABO28408	AbO28408 Human sec	500	44	54.3	23	7	ADN59796	Adn59796 Peptide-v
428	49	60.5	199	7	ABM23189	Abm23189 Human sec	501	44	54.3	25	7	ADN59726	Adn59726 Thrombopo
429	49	60.5	199	7	ABM30509	Abm30509 Human sec	502	44	54.3	25	7	ADN59728	Adn59728 Thrombopo
430	49	60.5	199	7	ABM21969	Abm21969 Human sec	503	44	54.3	35	7	ADN59754	Adn59754 Peptide-v
431	49	60.5	199	7	ABM21664	Abm21664 Human sec	504	44	54.3	36	7	ADN59766	Adn59766 Peptide-v
432	49	60.5	199	7	ABM15195	Abm15195 Human sec	505	44	54.3	36	7	ADN59756	Adn59756 Peptide-v
433	49	60.5	199	7	ABO41270	AbO41270 Human sec	506	44	54.3	36	7	ADN59767	Adn59767 Peptide-v
434	49	60.5	199	7	ABO37000	AbO37000 Human sec	507	44	54.3	41	7	ADN59816	Adn59816 Peptide-
435	49	60.5	199	7	ABO37610	AbO37610 Human sec	508	44	54.3	41	7	ADN59772	Adn59772 Peptide-v
436	49	60.5	199	7	ABM75400	Abm75400 Human sec	509	44	54.3	41	7	ADN59773	Adn59773 Peptide-v
437	49	60.5	199	7	ABM33680	Abm33680 Human sec	510	44	54.3	46	7	ADN59790	Adn59790 Peptide-v
438	49	60.5	199	7	ABO46435	AbO46435 Human PRO	511	44	54.3	46	7	ADN59785	Adn59785 Peptide-v
439	49	60.5	199	7	ADA83023	Ada83023 Human sec	512	44	54.3	46	7	ADN59784	Adn59784 Peptide-v
440	49	60.5	199	7	ABM32034	Abm32034 Human sec	513	44	54.3	46	7	ADN59791	Adn59791 Peptide-v
441	49	60.5	199	7	ABM31424	Abm31424 Human sec	514	44	54.3	75	7	ADN59758	Adn59758 Peptide-v
442	49	60.5	199	7	ADB86331	AdB86331 Human sec	515	44	54.3	148	8	ADI42202	Adi42202 Plant tra
443	49	60.5	199	7	ABM32339	Abm32339 Human sec	516	44	54.3	148	8	ADO02695	Ado02695 Thalecres
444	49	60.5	199	7	ABM32644	Abm32644 Human sec	517	44	54.3	148	8	ADO62315	Ado62315 Transcrip
445	49	60.5	199	7	ABM31729	Abm31729 Human sec	518	44	54.3	1039	7	ADJ70453	Adj70453 Human hea
446	49	60.5	199	7	ABM31119	Abm31119 Human sec	519	44	54.3	1098	7	ADE60625	Ade60625 Rat Prote
447	49	60.5	199	7	ADD06061	AdD06061 Human hum	520	44	54.3	1126	6	ABU70708	Abu70708 Human adi
449	49	60.5	199	7	ADG03056	AdG03056 Novel hum	521	44	54.3	1126	8	ABM80766	Abm80766 Tumour-as
448	49	60.5	199	7	ADG01763	AdG01763 Novel hum	522	44	54.3	1126	8	ABM80767	Abm80767 Tumour-as
450	49	60.5	199	7	ADF95938	AdF95938 Novel hum	523	44	54.3	1132	7	ADE60627	Ade60627 Human Pro
451	49	60.5	199	7	ADG12753	AdG12753 Novel hum	524	44	54.3	1132	7	ADE60623	Ade60623 Human Pro
452	49	60.5	199	7	ADH09413	AdH09413 Human PRO	525	44	54.3	1132	8	ABM80765	Abm80765 Tumour-as
453	49	60.5	199	7	ADL33192	AdL33192 Novel hum	526	43	53.1	18	7	ADN59678	Adn59678 Thrombopo
454	49	60.5	199	7	ADM30728	Adm30728 Novel hum	527	43	53.1	18	7	ADN59677	Adn59677 Thrombopo
455	49	60.5	199	8	ADE74725	Ade74725 Human sec	528	43	53.1	25	7	ADN59734	Adn59734 Thrombopo
456	49	60.5	199	8	ADE75337	Ade75337 Human sec	529	43	53.1	25	7	ADN59736	Adn59736 Thrombopo
457	49	60.5	199	8	ADF96550	AdF96550 Novel hum	530	43	53.1	62	3	AAB42889	Aab42889 Human ORF
458	49	60.5	199	8	ADG04821	AdG04821 Novel hum	531	42	51.9	14	4	AAM98500	Aam98500 Human pep
459	49	60.5	199	8	ADG00981	AdG00981 Novel hum	532	42	51.9	29	2	AAO28894	Aay028894 Fragment
460	49	60.5	199	8	ADG83237	AdG83237 Human PRO	533	42	51.9	29	7	ADA07724	Ada07724 Human sec
461	49	60.5	199	8	ADH26518	AdH26518 Novel hum	534	42	51.9	29	8	ADN41333	Adn41333 Novel hum
462	49	60.5	199	8	ADH33487	Adh33487 Human PRO	535	42	51.9	43	3	AAG57089	Aeg57089 Arabidops

536	42	51.9	93	4	AAU50884	Aau50884 Propionib	609	41	50.6	126	8	ADO10073	Ado10073 Novel hum
537	42	51.9	93	6	ABM47403	Abm47403 Propionib	610	41	50.6	159	8	ADU02995	Adu02995 Novel hum
538	42	51.9	100	2	AAY59831	Aay59831 Human nor	611	41	50.6	173	8	ADU02996	Adu02996 Novel hum
539	42	51.9	143	8	ADL04815	Adl04815 M. catarr	612	41	50.6	203	6	ABP99502	Abp99502 Human sec
540	42	51.9	171	8	ABO58540	Abo58540 Human gen	613	41	50.6	203	6	ABR00975	Abr00975 Human gen
541	42	51.9	171	8	ABO58663	Abo58663 Human gen	614	41	50.6	203	6	ADA98177	Ada98177 Human sec
542	42	51.9	189	5	AAO17523	Aao17523 Human M32	615	41	50.6	203	6	ADA44031	Ada44031 Human sec
543	42	51.9	196	7	ADC32882	Adc32882 Human nov	616	41	50.6	203	7	ADC20347	Adc20347 Human sec
544	42	51.9	216	3	AAB12489	Aab12489 Human PSE	617	41	50.6	203	7	ADF10687	Adf10687 Human sec
545	42	51.9	225	4	AAU25636	Aau25636 G protein	618	41	50.6	204	3	AAB51840	Aab51840 Human sec
546	42	51.9	300	7	ADC86753	Adc86753 Human GPC	619	41	50.6	205	3	AAB32114	Aab32114 Human sec
547	42	51.9	419	5	ABB07921	Abb07921 Murine pe	620	41	50.6	320	4	AAB94502	Aab94502 Human pro
548	42	51.9	419	8	ADP48669	Adp48669 Mouse Pel	621	41	50.6	320	4	AAG67255	Aag67255 Amino aci
549	42	51.9	420	5	ABB78077	Abb78077 Amino aci	622	41	50.6	320	5	AAO17516	Aao17516 Human M30
550	42	51.9	420	5	ABB07922	Abb07922 Human pel	623	41	50.6	320	5	AAO17517	Aao17517 Human M30
551	42	51.9	420	5	AAO17522	Aao17522 Human M33	624	41	50.6	418	5	ABB78078	Abb78078 Amino aci
552	42	51.9	420	6	ABR41180	Abrc41180 Human nov	625	41	50.6	418	5	ABB07919	Abb07919 Murine pe
553	42	51.9	420	7	ADC31180	Adc31180 Human pel	626	41	50.6	418	5	ABB07920	Abb07920 Human pel
554	42	51.9	420	8	ADP48671	Adp48671 Human Pel	627	41	50.6	418	5	AAO17515	Aao17515 Rat M30 p
555	42	51.9	428	4	AAB95188	Aab95188 Human pro	628	41	50.6	418	5	AAO17519	Aao17519 Human M30
556	42	51.9	487	4	AAB88569	Aab88569 Human hyd	629	41	50.6	418	5	AAO17518	Aao17518 Human M30
557	42	51.9	487	6	ABJ20214	Abj20214 Human SLC	630	41	50.6	418	6	ABU00017	Abu00017 Human nov
558	42	51.9	487	6	ABJ20215	Abj20215 Human SLC	631	41	50.6	418	8	ADF83103	Adf83103 Human Pel
559	42	51.9	487	6	ABU04694	Abu04694 Human exp	632	41	50.6	418	8	ADO44005	Ado44005 Amino aci
560	42	51.9	487	6	ABU04693	Abu04693 Human exp	633	41	50.6	418	8	ADP48665	Adp48665 Mouse Pel
561	42	51.9	487	6	ABU04696	Abu04696 Human exp	634	41	50.6	418	8	ADP48667	Adp48667 Human Pel
562	42	51.9	487	6	ABU04699	Abu04699 Human exp	635	41	50.6	418	8	ADP56420	Adp56420 Human PRO
563	42	51.9	487	6	ABU04700	Abu04700 Human exp	636	41	50.6	418	8	ADP24929	Adp24929 PRO polyp
564	42	51.9	487	6	ABU04698	Abu04698 Human exp	637	41	50.6	418	9	ADY20614	Ady20614 PRO polyp
565	42	51.9	487	6	ABU04697	Abu04697 Human exp	638	41	50.6	418	9	ADY17390	Ady17390 PRO polyp
566	42	51.9	487	6	ABU04695	Abu04695 Human exp	639	41	50.6	418	9	ADA27314	Ada27314 Rat BAT1
567	42	51.9	487	8	ADM68409	Adm68409 Human sol	640	41	50.6	504	4	AAG89259	Aag89259 Human sec
568	42	51.9	497	2	AAW37734	Aaw37734 Human cyt	641	41	50.6	504	6	AAE36234	Aae36234 Human arg
569	42	51.9	497	2	AAW44160	Aaw44160 Human ret	642	41	50.6	586	2	AAAY23642	Aay23642 Protein e
570	42	51.9	497	4	AAB85154	Aab85154 Cytochrom	643	41	50.6	840	6	AAAY23642	Aay23642 Protein e
571	42	51.9	497	5	ABP52145	Abp52145 Human P45	644	41	50.6	840	6	ABU54642	Abu54642 Human NOV
572	42	51.9	497	5	AAE15326	Aae15326 Human P45	645	41	50.6	840	8	ADL30844	Adl30844 Human pro
573	42	51.9	497	6	AAE35260	Aae35260 Human P45	646	41	50.6	878	4	AAAB93303	Aab93303 Human pro
574	42	51.9	497	8	ADU66825	Adu66825 Human RA-	647	41	50.6	878	4	ADJ96614	Adj96614 Human mic
575	42	51.9	497	9	ADV90765	Adv90765 Human ret	648	41	50.6	878	8	ADM72229	Adm72229 Human TAS
576	42	51.9	497	9	ADY62582	Ady62582 Human P45	649	41	50.6	878	8	ADQ89836	Adq89836 Antagonis
577	42	51.9	517	6	ABB82492	Abb82492 A. aegypt	650	41	50.6	879	4	AAG66515	Aag66515 Human 184
578	42	51.9	539	3	AAV58044	Aay58044 Arabidops	651	41	50.6	879	6	ABU54641	Abu54641 Human NOV
579	42	51.9	877	6	ABP96076	Abp96076 Human pro	652	41	50.6	1238	6	AAE36233	Aae36233 Human arg
580	41.5	51.2	18	7	ADN59681	Adn59681 Thrombopo	653	41	50.6	1584	4	ABG04239	Abg04239 Novel hum
581	41.5	51.2	25	7	ADN59742	Adn59742 Thrombopo	654	40	49.4	14	2	AAW09479	Aaw09479 Thrombopo
582	41.5	51.2	25	7	ADN59752	Adn59752 Peptide-v	655	40	49.4	14	2	AAW09485	Aaw09485 Thrombopo
583	41	50.6	13	3	AAAB17015	Aab17015 TPO-mimet	656	40	49.4	14	2	AAW36630	Aaw36630 Thrombopo
584	41	50.6	13	5	ABB72901	Abb72901 TPO mimet	657	40	49.4	14	2	AAW36636	Aaw36636 Thrombopo
585	41	50.6	13	7	ADJ73054	Adj73054 TPO mimet	658	40	49.4	14	2	AAW36631	Aaw36631 Thrombopo
586	41	50.6	13	7	ADJ73052	Adj73052 TPO mimet	659	40	49.4	14	3	AAAB17013	Aab17013 TPO-mimet
587	41	50.6	13	7	ADJ73056	Adj73056 TPO mimet	660	40	49.4	14	4	AAU25849	Aau25849 Human thr
588	41	50.6	13	7	ADJ73053	Adj73053 TPO mimet	661	40	49.4	14	4	AAU25855	Aau25855 Human thr
589	41	50.6	13	7	ADJ73055	Adj73055 TPO mimet	662	40	49.4	14	4	AAU25850	Aau25850 Human thr
590	41	50.6	13	8	ADJ52687	Adj52687 CH1 delet	663	40	49.4	14	5	ABB72899	Abb72899 TPO mimet
591	41	50.6	13	8	ADJ51648	Adj51648 CH1 delet	664	40	49.4	14	7	ADJ73050	Adj73050 TPO mimet
592	41	50.6	18	7	ADN59680	Adn59680 Thrombopo	665	40	49.4	14	8	ADJ52685	Adj52685 CH1 delet
593	41	50.6	18	7	ADN59671	Adn59671 Thrombopo	666	40	49.4	18	7	ADN59676	Adn59676 Thrombopo
594	41	50.6	22	7	ADN59838	Adn59838 TMP pepti	667	40	49.4	25	7	ADN59732	Adn59732 Thrombopo
595	41	50.6	23	7	ADN59795	Adn59795 Peptide-v	668	40	49.4	62	2	AAAY11422	Aay11422 Human 5'
596	41	50.6	23	7	ADN59777	Adn59777 Peptide-v	669	40	49.4	79	4	AAW33730	Aam33730 Peptide #
597	41	50.6	25	7	ADN59724	Adn59724 Thrombopo	670	40	49.4	79	4	AAW73536	Aam73536 Human bon
598	41	50.6	25	7	ADN59740	Adn59740 Thrombopo	671	40	49.4	79	4	ABG55263	Abg55263 Human liv
599	41	50.6	36	7	ADN59765	Adn59765 Peptide-v	672	40	49.4	79	4	ABG43399	Abg43399 Human pep
600	41	50.6	41	7	ADN59771	Adn59771 Peptide-v	673	40	49.4	109	7	ADD95172	Add95172 Bacteriop
601	41	50.6	46	7	ADN59789	Adn59789 Peptide-v	674	40	49.4	112	6	ABP75402	Abp75402 Human sec
602	41	50.6	46	7	ADN59757	Adn59757 Peptide-v	675	40	49.4	118	5	AAE14256	Aae14256 Human 231
603	41	50.6	46	7	ADN59783	Adn59783 Peptide-v	676	40	49.4	123	5	AAW48387	Aam48387 Human YSK
604	41	50.6	77	1	AAP30661	Aap30661 Enkephali	677	40	49.4	155	8	ADX75528	Adx75528 Plant ful
605	41	50.6	84	4	AAU41991	Aau41991 Propionib	678	40	49.4	168	7	ADG28434	Adg28434 Soybean S
606	41	50.6	84	6	ABM38510	Abm38510 Propionib	679	40	49.4	320	4	AAW93893	Aam93893 Human pol
607	41	50.6	104	5	ABB89467	Abb89467 Human pol	680	40	49.4	320	8	ADL31994	Adl31994 Human pro
608	41	50.6	113	1	AAP30659	Aap30659 Enkephali	681	40	49.4	320			

682	40	49.4	406	7	ADB65774	Adb65774	Human pro	755	38	46.9	105	6	ABU55355	Abu55355	Human nov
683	40	49.4	409	2	AAW26526	Aaw26526	Branched	756	38	46.9	141	6	ABM72249	Abm72249	Staphyloc
684	40	49.4	409	4	AAB61179	Aab61179	Pseudomon	757	38	46.9	151	4	AAU44520	Aau44520	Propionib
685	40	49.4	410	6	ABU40351	Abu40351	Protein e	758	38	46.9	151	6	ABM41039	Abm41039	Propionib
686	40	49.4	415	8	ADO44009	Ado44009	Amino aci	759	38	46.9	154	8	ADQ66223	Adq66223	Novel hum
687	40	49.4	419	6	ABR41082	Abr41082	Mouse MAP	760	38	46.9	158	4	AAM39969	Aam39969	Human pol
688	40	49.4	419	8	ADQ67161	Adq67161	Novel hum	761	38	46.9	166	2	AAW94848	Aaw94848	Human cer
689	40	49.4	445	5	ABB07923	Abb07923	Human pel	762	38	46.9	166	8	ADR08617	Adr08617	Human pro
690	40	49.4	445	5	AAO17520	Aao17520	Murine M3	763	38	46.9	176	8	ADX93119	Adx93119	Plant ful
691	40	49.4	445	5	AAO17521	Aao17521	Human M31	764	38	46.9	177	4	AAE03109	Aae03109	Human gen
692	40	49.4	445	7	ADC31807	Adc31807	Human nov	765	38	46.9	177	5	ABG63620	Abg63620	Human alb
693	40	49.4	445	7	ADD71147	Add71147	Human int	766	38	46.9	177	8	ADL76885	Adl76885	Albumin f
694	40	49.4	445	7	ADP48675	Adp48675	Human pel	767	38	46.9	180	6	ABP96348	Abp96348	Human AGE
695	40	49.4	464	7	ABO67914	Abo67914	Pseudomon	768	38	46.9	181	9	ABM95454	Abm95454	M. xanthu
696	40	49.4	476	8	ADS42549	Ads42549	Bacterial	769	38	46.9	186	8	ADX97029	Adx97029	Plant ful
697	40	49.4	506	6	ADA55185	Ada55185	Human pro	770	38	46.9	206	4	AAM41755	Aam41755	Human pol
698	40	49.4	554	7	ADM04654	Adm04654	Human pro	771	38	46.9	206	9	ADX08152	Adx08152	Cyclin-de
699	40	49.4	561	8	ADN27155	Adn27155	Bacterial	772	38	46.9	219	2	AAW24021	Aaw24021	Human pho
700	40	49.4	715	9	ABM93740	Abm93740	M. xanthu	773	38	46.9	219	9	ADX08150	Adx08150	Cyclin-de
701	40	49.4	881	5	ADH48730	Adh48730	NOV6 prot	774	38	46.9	235	8	ADX75054	Adx75054	Plant ful
702	40	49.4	881	8	ADU02359	Adu02359	Novel hum	775	38	46.9	264	4	AAG71897	Aag71897	Human olf
703	40	49.4	1068	7	ADC39124	Adc39124	Novel hum	776	38	46.9	267	2	AAR91362	Aar91362	Proenkeph
704	40	49.4	1077	4	AAU01924	Aau01924	Human ade	777	38	46.9	267	4	AAG89159	Aag89159	Human sec
705	40	49.4	1077	8	ADL14164	Adl14164	Novel hum	778	38	46.9	267	8	ABG75474	Abg75474	Human bet
706	40	49.4	1086	5	AAE17129	Aae17129	Human ade	779	38	46.9	267	8	ABG75473	Abg75473	Human pro
707	40	49.4	1103	5	ABB81465	Abb81465	Human ade	780	38	46.9	289	5	AAU95669	Aau95669	Human olf
708	39	48.1	58	8	ABO56054	Abo56054	Human gen	781	38	46.9	289	8	ADG83416	Adg83416	Human Olf
709	39	48.1	76	4	AAU61033	Aau61033	Propionib	782	38	46.9	297	3	AAG52752	Aag52752	Arabidops
710	39	48.1	76	6	ABM57552	Abm57552	Propionib	783	38	46.9	297	3	AAG16893	Aag16893	Arabidops
711	39	48.1	77	4	AAU53862	Aau53862	Propionib	784	38	46.9	297	4	ABB11849	Abb11849	Human pre
712	39	48.1	77	6	ABM50381	Abm50381	Propionib	785	38	46.9	313	6	ABR01653	Abr01653	Human G p
713	39	48.1	82	6	ABP58654	Abp58654	Human mil	786	38	46.9	319	6	ABR01638	Abr01638	Human G p
714	39	48.1	260	4	AAW25622	Aam25622	Human pro	787	38	46.9	320	5	ABG76863	Abg76863	Human G-p
715	39	48.1	260	6	ABO00904	Abo00904	Polypepti	788	38	46.9	320	5	ABP95882	Abp95882	Human GPC
716	39	48.1	308	5	AAU95592	Aau95592	Human olf	789	38	46.9	320	5	AAU95784	Aau95784	Human olf
717	39	48.1	308	6	ABO00671	Abo00671	Novel hum	790	38	46.9	320	5	AAU95782	Aau95782	Human G p
718	39	48.1	324	4	AAG71587	Aag71587	Human olf	791	38	46.9	320	6	ABP97086	Abp97086	Human G p
719	39	48.1	324	5	ABP51609	Abp51609	Human G p	792	38	46.9	320	6	ABP97085	Abp97085	Human G p
720	39	48.1	324	7	ADI04343	Adi04343	Human G-p	793	38	46.9	320	7	ADC86249	Adc86249	Human GPC
721	39	48.1	335	4	AAU24741	Aau24741	Human olf	794	38	46.9	320	7	ADC86247	Adc86247	Human GPC
722	39	48.1	335	5	ABB06604	Abb06604	G protein	795	38	46.9	320	8	ADG83532	Adg83532	Human Olf
723	39	48.1	335	5	ABB06605	Abb06605	G protein	796	38	46.9	320	8	ADG83536	Adg83536	Human Olf
724	39	48.1	335	5	ABB06606	Abb06606	G protein	797	38	46.9	333	3	AAG16892	Aag16892	Arabidops
725	39	48.1	335	5	ABP95761	Abp95761	Human GPC	798	38	46.9	333	3	AAG52751	Aag52751	Arabidops
726	39	48.1	335	5	AAU85361	Aau85361	G-coupled	799	38	46.9	337	3	AAG52750	Aag52750	Arabidops
727	39	48.1	335	7	ADC85603	Adc85603	Human GPC	800	38	46.9	337	3	AAG16891	Aag16891	Arabidops
728	39	48.1	341	3	AAAY69880	Aay69880	B. lactof	801	38	46.9	368	4	ABG03113	Abg03113	Novel hum
729	39	48.1	379	8	ADX74528	Adx74528	Plant ful	802	38	46.9	426	7	ABO62034	Abo62034	Klebsiell
730	39	48.1	460	8	ADP79710	Adp79710	Cycloclas	803	38	46.9	433	4	ABG04452	Abg04452	Novel hum
731	39	48.1	461	9	ABM94829	Abm94829	Human TMD	804	38	46.9	433	7	ADE09953	Ade09953	Novel pro
732	39	48.1	491	7	ADG84110	Adg84110	Human TMD	805	38	46.9	433	9	AEA21042	Aea21042	Novel hum
733	39	48.1	517	4	ABB62049	Abb62049	Drosophil	806	38	46.9	435	8	ADY11835	Ady11835	Plant ful
734	39	48.1	716	6	ABU33863	Abu33863	Protein e	807	38	46.9	447	7	ADC31461	Adc31461	Human nov
735	39	48.1	966	6	ABU56300	Abu56300	Caspase r	808	38	46.9	450	4	AAU56552	Aau56552	Propionib
736	39	48.1	1086	6	ABU46534	Abu46534	Protein e	809	38	46.9	450	6	ABM53071	Abm53071	Propionib
737	39	48.1	1086	9	AEB91670	Aeb91670	Microbial	810	38	46.9	464	3	AAB53151	Aab53151	Macaca mu
738	39	48.1	1525	7	ABM89427	Abm89427	Rice abio	811	38	46.9	476	8	ADT58335	Adt58335	Plant pol
739	38.5	47.5	265	7	ABO63876	Abo63876	Klebsiell	812	38	46.9	493	5	ABB93470	Abb93470	Herbicide
740	38	46.9	22	8	ADS20872	Ads20872	TMIII-VII	813	38	46.9	524	5	ABU65112	Abu65112	Human NOV
741	38	46.9	36	4	ABB39297	Abb39297	Peptide #	814	38	46.9	524	8	ADN61875	Adn61875	Human nov
742	38	46.9	36	4	AAM32805	Aam32805	Peptide #	815	38	46.9	526	8	ADY07792	Ady07792	Plant ful
743	38	46.9	36	4	AAM72565	Aam72565	Human bon	816	38	46.9	530	5	ABU65113	Abu65113	Human NOV
744	38	46.9	36	4	AAM59965	Aam59965	Human bra	817	38	46.9	530	8	ADN61877	Adn61877	Human nov
745	38	46.9	36	4	ABG54257	Abg54257	Human liv	818	38	46.9	545	7	ABO80203	Abo80203	Pseudomon
746	38	46.9	36	5	ABG42386	Abg42386	Human pep	819	38	46.9	549	3	AAAY73346	Aay73346	HTRM clon
747	38	46.9	37	3	AAB34113	Aab34113	Human sec	820	38	46.9	577	4	ABG13760	Abg13760	Novel hum
748	38	46.9	84	5	AAE20819	Aae20819	Human gen	821	38	46.9	580	3	AAB43912	Aab43912	Human can
749	38	46.9	84	5	ABG64683	Abg64683	Human alb	822	38	46.9	580	7	ADF55450	Adf55450	Human nov
750	38	46.9	84	8	ADL77950	Adl77950	Albumin f	823	38	46.9	603	4	ABG00654	Abg00654	Novel hum
751	38	46.9	91	8	ADX92954	Adx92954	Plant ful	824	38	46.9	603	4	ABG08209	Abg08209	Novel hum
752	38	46.9	105	4	AAU16286	Aau16286	Human nov	825	38	46.9	739	4	ABG03917	Abg03917	Novel hum
753	38	46.9	105	4	AAU29502	Aau29502	Human G p	826	38	46.9	781	4	ABG24696	Abg24696	Novel hum
754	38	46.9	105	5	ABG60790	Abg60790	Novel G p	827	38	46.9	790	6	ADA55066	Ada55066	Human pro

828	38	46.9	790	7	ADC31848	Adc31848 Human nov	901	37	45.7	61	6	ABP76329	Abp76329 Human GEN
829	38	46.9	790	7	ADJ71094	Adj71094 Human hea	902	37	45.7	65	4	AAU63303	Aau63303 Propionib
830	38	46.9	855	8	ADN74169	Adn74169 Thale cre	903	37	45.7	65	6	ABM59822	Abm59822 Propionib
831	38	46.9	994	8	ADQ39393	Adq39393 Human myo	904	37	45.7	72	4	AAU49140	Aau49140 Propionib
832	38	46.9	994	8	ADQ39395	Adq39395 Human myo	905	37	45.7	72	6	ABM45659	Abm45659 Propionib
833	38	46.9	994	8	ADQ39396	Adq39396 Human myo	906	37	45.7	91	4	ABB15848	Abb15848 Human ner
834	38	46.9	1027	8	ADQ39394	Adq39394 Human myo	907	37	45.7	103	6	ADA55455	Ada55455 Human pro
835	38	46.9	1081	7	ADE40146	Ade40146 Human NOV	908	37	45.7	107	3	AAB41227	Aab41227 Human ORF
836	38	46.9	1116	5	AAU97582	Aau97582 Human CD1	909	37	45.7	107	5	ABP04416	Abp04416 Human ORF
837	38	46.9	1116	8	ADQ17792	Adq17792 Human sof	910	37	45.7	110	8	ADL92620	Adl92620 IMABIS003
838	38	46.9	1116	9	ADV70266	Adv70266 Tumor-ass	911	37	45.7	110	8	ADR41966	Adr41966 VAP relat
839	38	46.9	1121	4	AAM39493	Aam39493 Human pol	912	37	45.7	110	9	ADX70284	Adx70284 Versatile
840	38	46.9	1124	4	AAM41280	Aam41280 Human pol	913	37	45.7	111	4	ABG02805	Abg02805 Novel hum
841	38	46.9	1124	4	AAM41279	Aam41279 Human pol	914	37	45.7	123	2	AAY76599	Aay76599 Human ova
842	38	46.9	1149	5	AAU97583	Aau97583 Human CD1	915	37	45.7	140	7	ABO80854	Abo80854 Pseudomon
843	38	46.9	1151	5	AAU97585	Aau97585 Human CD1	916	37	45.7	143	4	AAM14876	Aam14876 Peptide #
844	38	46.9	1154	4	AAM39494	Aam39494 Human pol	917	37	45.7	143	4	AAM14879	Aam14879 Peptide #
845	38	46.9	1156	5	AAU97584	Aau97584 Human CD1	918	37	45.7	143	4	ABB33848	Abb33848 Peptide #
846	38	46.9	1156	8	ADQ39397	Adq39397 Human myo	919	37	45.7	143	4	ABB33845	Abb33845 Peptide #
847	38	46.9	1232	8	ADE45189	Ade45189 Human CD1	920	37	45.7	143	4	AAM27305	Aam27305 Peptide #
848	38	46.9	1492	6	AAO31008	Aao31008 Human tra	921	37	45.7	143	4	AAM27308	Aam27308 Peptide #
849	38	46.9	1492	7	ADM29378	Adm29378 Human nov	922	37	45.7	143	4	ABB28663	Abb28663 Peptide #
850	38	46.9	1492	8	ABM82750	Abm82750 Human dia	923	37	45.7	143	4	ABB28661	Abb28661 Peptide #
851	38	46.9	1545	2	AAW33362	Aaw33362 Human can	924	37	45.7	143	4	ABB19289	Abb19289 Protein #
852	38	46.9	1545	2	AAW55966	Aaw55966 Human can	925	37	45.7	143	4	ABB19287	Abb19287 Protein #
853	38	46.9	1545	5	AAE29040	Aae29040 Homo cmo	926	37	45.7	143	4	AAM67018	Aam67018 Human bon
854	38	46.9	1545	5	ABP52109	Abp52109 Homo sapi	927	37	45.7	143	4	AAM67016	Aam67016 Human bon
855	38	46.9	1545	8	ADL60256	Adl60256 Human org	928	37	45.7	143	4	AAM54610	Aam54610 Human bra
856	38	46.9	1545	8	ADL60270	Adl60270 Human org	929	37	45.7	143	4	AAM54612	Aam54612 Human bra
857	38	46.9	1545	8	ADL60264	Adl60264 Human org	930	37	45.7	143	4	ABG48681	Abg48681 Human liv
858	38	46.9	1545	8	ADL60272	Adl60272 Human org	931	37	45.7	143	4	ABG48683	Abg48683 Human liv
859	38	46.9	1545	8	ADL60268	Adl60268 Human org	932	37	45.7	143	4	AAM02603	Aam02603 Peptide #
860	38	46.9	1545	8	ADL60250	Adl60250 Human org	933	37	45.7	143	4	AAM02601	Aam02601 Peptide #
861	38	46.9	1545	8	ADL60258	Adl60258 Human org	934	37	45.7	143	5	ABG36675	Abg36675 Human pep
862	38	46.9	1545	8	ADL60248	Adl60248 Human org	935	37	45.7	143	5	ABG36673	Abg36673 Human pep
863	38	46.9	1545	8	ADL60828	Adl60828 Human org	936	37	45.7	143	8	ABO59053	Abo59053 Human gen
864	38	46.9	1545	8	ADL60228	Adl60228 Human org	937	37	45.7	147	8	ADR09851	Adr09851 Human pro
865	38	46.9	1545	8	ADL60252	Adl60252 Human org	938	37	45.7	152	3	AAG59906	Aag59906 Arabidops
866	38	46.9	1545	8	ADL60254	Adl60254 Human org	939	37	45.7	177	4	ABG28550	Abg28550 Novel hum
867	38	46.9	1545	8	ADL60260	Adl60260 Human org	940	37	45.7	205	4	ABG02806	Abg02806 Novel hum
868	38	46.9	1545	8	ADL60262	Adl60262 Human org	941	37	45.7	222	4	ABG10784	Abg10784 Novel hum
869	38	46.9	1545	8	ADL60266	Adl60266 Human org	942	37	45.7	248	5	AAU91164	Aau91164 Pantothien
870	38	46.9	1545	8	ADP20017	Adp20017 Human ATP	943	37	45.7	249	8	ADS21323	Ads21323 Bacterial
871	37.5	46.3	198	9	ADZ13371	Adz13371 Human can	944	37	45.7	252	3	AAG06656	Aag06656 Arabidops
872	37.5	46.3	509	3	AAB13558	Aab13558 Streptomy	945	37	45.7	253	7	ABO72591	Abo72591 Pseudomon
873	37	45.7	15	3	AAB17018	Aab17018 TPO-mimet	946	37	45.7	266	7	ADC85016	Adc85016 HLA DRB1
874	37	45.7	15	5	ABB72904	Abb72904 TPO mimet	947	37	45.7	266	7	ADC85013	Adc85013 HLA DRB1
875	37	45.7	15	8	ADJ52691	Adj52691 CH1 delet	948	37	45.7	266	7	ADC85019	Adc85019 HLA DRB1
876	37	45.7	15	8	ADJ52690	Adj52690 CH1 delet	949	37	45.7	267	8	ABO60229	Abo60229 Human gen
877	37	45.7	15	8	ADJ51652	Adj51652 CH1 delet	950	37	45.7	272	3	AAG06655	Aag06655 Arabidops
878	37	45.7	15	8	ADJ51651	Adj51651 CH1 delet	951	37	45.7	281	3	AAG06654	Aag06654 Arabidops
879	37	45.7	18	2	AAW09460	Aaw09460 Thrombopo	952	37	45.7	282	3	AAG60574	Aag60574 Arabidops
880	37	45.7	18	2	AAW09498	Aaw09498 Thrombopo	953	37	45.7	282	3	AAG60570	Aag60570 Arabidops
881	37	45.7	18	2	AAW36649	Aaw36649 Thrombopo	954	37	45.7	282	3	AAG20635	Aag20635 Arabidops
882	37	45.7	18	2	AAW33027	Aaw33027 Thrombopo	955	37	45.7	284	5	ABB93853	Abb93853 Herbicida
883	37	45.7	18	2	AAW36652	Aaw36652 Thrombopo	956	37	45.7	289	4	ABG09868	Abg09868 Novel hum
884	37	45.7	18	3	AAB17026	Aab17026 TPO-mimet	957	37	45.7	310	4	AAG71915	Aag71915 Human olf
885	37	45.7	18	4	AAU25868	Aau25868 Human thr	958	37	45.7	310	4	AAU24739	Aau24739 Human olf
886	37	45.7	18	4	AAU25824	Aau25824 Human thr	959	37	45.7	310	5	ABG76839	Abg76839 Human G-p
887	37	45.7	18	4	AAU25871	Aau25871 Human thr	960	37	45.7	310	5	ABP51597	Abp51597 Human G p
888	37	45.7	18	5	ABB72912	Abb72912 TPO mimet	961	37	45.7	310	5	ABP95708	Abp95708 Human GPC
889	37	45.7	18	7	ADJ73064	Adj73064 TPO mimet	962	37	45.7	310	5	AAU95691	Aau95691 Human olf
890	37	45.7	18	7	ADN59657	Adn59657 Thrombopo	963	37	45.7	310	5	AAU85359	Aau85359 G-coupled
891	37	45.7	18	8	ADJ52699	Adj52699 CH1 delet	964	37	45.7	310	6	ABU11155	Abu11155 Human G-p
892	37	45.7	18	8	ADJ51660	Adj51660 CH1 delet	965	37	45.7	310	7	ADC86123	Adc86123 Human GPC
893	37	45.7	22	7	ADN59824	Adn59824 TMP pepti	966	37	45.7	310	7	ADI04319	Adi04319 Human G-p
894	37	45.7	23	7	ADN59775	Adn59775 Peptide-v	967	37	45.7	319	3	AAU90513	Aau90513 Rat cereb
895	37	45.7	23	7	ADN59793	Adn59793 Peptide-v	968	37	45.7	319	7	ADD29436	Add29436 Rat G-pro
896	37	45.7	25	7	ADN59696	Adn59696 Thrombopo	969	37	45.7	319	8	ABO58539	Abo58539 Human gen
897	37	45.7	36	7	ADN59763	Adn59763 Peptide-v	970	37	45.7	344	2	AAW30208	Aaw30208 C. albica
898	37	45.7	41	7	ADN59769	Adn59769 Peptide-v	971	37	45.7	344	2	AAW35372	Aaw35372 C. albica
899	37	45.7	46	7	ADN59781	Adn59781 Peptide-v	972	37	45.7	344	5	ABP73737	Abp73737 Candida a
900	37	45.7	46	7	ADN59787	Adn59787 Peptide-v	973	37	45.7	375	7	ADC94127	Adc94127 E. faeciu

974 37 45.7 410 6 ABU22259
975 37 45.7 410 6 ABU20206
976 37 45.7 410 9 AEB38098
977 37 45.7 410 9 AEB41389
978 37 45.7 416 4 AAU04813
979 37 45.7 416 6 ABP99322
980 37 45.7 422 7 ABO60778
981 37 45.7 427 9 AEB42134
982 37 45.7 431 9 AEB38990
983 37 45.7 435 4 AAM39070
984 37 45.7 445 3 AAG60569
985 37 45.7 449 7 ABO68161
986 37 45.7 451 3 AAG60573
987 37 45.7 451 3 AAG20634
988 37 45.7 452 3 AAB29474
989 37 45.7 454 2 AAR97842
990 37 45.7 454 2 AAR93616
991 37 45.7 457 7 ADM04739
992 37 45.7 458 9 AEB28955
993 37 45.7 483 4 AAM40856
994 37 45.7 504 5 ABB89686
995 37 45.7 513 4 ABB65614
996 37 45.7 538 4 ABG21068
997 37 45.7 566 4 ABB61743
998 37 45.7 575 5 ABP27731
999 37 45.7 625 5 ABB81615
1000 37 45.7 636 4 AAE05105

Abu22259 Protein e
Abu20206 Protein e
Aeb38098 L. pneumo
Aeb41389 L. pneumo
Aau04813 Micromono
Abp99322 Orthosomy
Abo60778 Klebsiell
Aeb42134 L. pneumo
Aeb38990 L. pneumo
Aam39070 Human pol
Aag60569 Arabidops
Abo68161 Pseudomon
Aag60573 Arabidops
Aag20634 Arabidops
Aab29474 Burkholde
Aar97842 Kaposi's
Aar93616 Kaposi's
Adm04739 Human pro
Aeb28955 Anopheles
Aam40856 Human pol
Abb89686 Human pol
Abb65614 Drosophil
Abg21068 Novel hum
Abb61743 Drosophil
Abp27731 Streptoco
Abb81615 Human HDA
Aae05105 Drosophil

ALIGNMENTS

RESULT 1
AAW09483
ID AAW09483 standard; protein; 14 AA.

AC AAW09483;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWLTC 14
Db 1 CELVGPSLMSWLTC 14

RESULT 2
AAW36634
ID AAW36634 standard; peptide; 14 AA.

AC AAW36634;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.

XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.

XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX DE Human thrombopoietin receptor (TPO-R) activator peptide #39.

XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

KW KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

KW KW bone marrow transplantation; haematological disorder; platelet disorder;

KW KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;

KW KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;

XX KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

OS KW Homo sapiens.

XX KW US6251864-B1.

PD KW 26-JUN-2001.

XX KW 01-MAR-2000; 2000US-00516704.

PF KW 07-JUN-1995; 95US-00478128.

PR KW 07-JUN-1995; 95US-00485301.

PR KW 07-JUN-1996; 96WO-US009623.

PR KW 15-AUG-1996; 96US-00699027.

XX KW (GLAX) GLAXO GROUP LTD.

PA KW Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

XX KW Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI KW Yin Q;

XX KW WPI; 2001-564142/63.

DR KW

XX KW Activating thrombopoietin receptors in cells, used to treat

PT KW thrombocytopenia and hematological disorders, comprises contacting cells

PT KW with peptides and peptide mimetics attached to hydrophilic polymers.

XX KW

PS KW Disclosure; Col 20; 128pp; English.

XX KW

CC KW Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

CC KW bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC KW of activating thrombopoietin receptors in cells comprise contacting the

CC KW cells with effective amounts of peptides and peptide mimetics attached to

CC KW hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC KW as that due to chemotherapy, radiation therapy or bone-marrow

CC KW transplantation and to prevent thrombocytopenia in patients at risk.The

CC KW sequences are used to treat and prevent haematological disorders

CC KW including thrombocytopenia and platelet disorders. They are used in vitro

CC KW as unique tools for understanding the biological role of thrombopoietin

CC KW (TPO) and to develop other compounds that bind to and activate the TPO

CC KW receptor. The peptides can be used to detect TPO receptors on living

CC KW cells and fixed cells, in biological fluids, in tissue homogenates, and

CC KW in purified or natural biological materials. They may also be used for in

CC KW situ staining, fluorescence-activated cell sorting, Western blotting and

CC KW enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC KW be used for in vitro expansion of megakaryocytes and their committed

CC KW progenitors alone or in conjunction with additional cytokines

XX KW

SQ KW Sequence 14 AA;

Query Match 100.0%; Score 81; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.5e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLTC 14

Db 1 CELVGPSLMSWLTC 14

RESULT 6

ABB72897

ID ABB72897 standard; peptide; 14 AA.

XX

AC ABB72897;

XX

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:67.

DE

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;

KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;

KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

XX KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;

XX KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;

PN KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

PR KW sleep disorder; neurological degenerative disease; anaemia;

XX KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;

XX KW Fanconi's syndrome.

OS KW Homo sapiens.

XX KW Synthetic.

PN WO200183525-A2.

XX KW 08-NOV-2001.

PD

XX KW 02-MAY-2001; 2001WO-US014310.

PF

XX KW 03-MAY-2000; 2000US-00563286.

PR

XX KW (AMGE-) AMGEN INC.

PA

XX KW Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

PI WPI; 2002-130313/17.

XX

DR Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX

PS Claim 39; Page 44; 176pp; English.

XX

CC The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 5; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.5e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLTC 14

Db 1 CELVGPSLMSWLTC 14

Db || ||:|: || |
 1 CEQDGPTLLEWLKC 14

RESULT 13
ABB72896
ID ABB72896 standard; peptide; 14 AA.
XX
AC ABB72896;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:66.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 44; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX

SQ Sequence 14 AA;
 Query Match 64.2%; Score 52; DB 5; Length 14;
 Best Local Similarity 57.1%; Pred. No. 0.12;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGP S L M S W L T C 14
 ||| ||:|: || |
Db 1 CEQDGPTLLEWLKC 14

RESULT 14
ADJ73047
ID ADJ73047 standard; peptide; 14 AA.
XX
AC ADJ73047;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 501.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 501; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 14 AA;
 Query Match 64.2%; Score 52; DB 7; Length 14;
 Best Local Similarity 57.1%; Pred. No. 0.12;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGP S L M S W L T C 14

Db 1 CEQDGTLLLEWLKC 14

RESULT 15

ADJ52682

ID ADJ52682 standard; peptide; 14 AA.

XX

AC ADJ52682;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID501.

XX

KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082870/08.

XX

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for

PT modulating, treating, alleviating, preventing an immune, cardiovascular,

PT or neurodegenerative disease or disorder, anemia, cancer, or infectious

PT diseases.

XX

PS Claim 2; SEQ ID NO 501; 129pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an immunosuppressive,

CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,

CC antibacterial, virucide or fungicide activity. In addition, the disclosed

CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody

CC is useful for diagnosing or treating a disease condition in a cell,

CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an

CC immune, cardiovascular (for example arrhythmia, hypertension or heart

CC failure), or neurodegenerative (for example multiple sclerosis, dementia

CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or

CC fungal infection). The present sequence is that of a peptide which may be

CC used during the creation of a mimetibody of the invention.

XX

SQ Sequence 14 AA;

Query Match 64.2%; Score 52; DB 8; Length 14;

Best Local Similarity 57.1%; Pred. No. 0.12;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWLTC 14

Db 1 CEQDGTLLLEWLKC 14

RESULT 16

ADJ51643

ID ADJ51643 standard; peptide; 14 AA.

XX

AC ADJ51643;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID501.

XX

KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; infectious disorder; immunological disorder;

KW allergic disorder; muscular disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002424-A2.

XX

PD 08-JAN-2004.

XX

PF 30-JUN-2003; 2003WO-US020495.

XX

PR 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX

DR WPI; 2004-082872/08.

XX

PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for

PT diagnosing, preventing or treating cardiovascular, dermatologic,

PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX

PS Claim 15; SEQ ID NO 501; 123pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CH1 deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,

CC obstetric, haematologic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX

SQ Sequence 14 AA;

Query Match 64.2%; Score 52; DB 8; Length 14;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGP S LMSW L T C 14
|| || |:|: || |
Db 1 CEQDGP T LLEW L K C 14

RESULT 17
ABB09517
ID ABB09517 standard; protein; 129 AA.
XX ABB09517;
AC ABB09517;
XX
DT 01-NOV-2002 (first entry)
XX
DE Human RNase-like NOV8 protein, SEQ ID NO:32.
XX
KW Human; NOVX; neurological disorder; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;
KW addition; tuberosus sclerosis; cancer; immune disorder; allergy;
KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;
KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;
KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;
KW pancreatic kidney disease; glomerular endothe liosis; bacterial infection;
KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;
KW atherosclerosis; cell signal processing-related disorder;
KW metabolic pathway regulation disorder; cytostatic; neuroprotective;
KW antiinflammatory; immunosuppressive; analgesic; antiatherosclerotic;
KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;
KW differentiation; proliferation; motility; haematopoiesis; wound healing;
KW angiogenesis; forensic biology; transgenic animal; drug screening;
KW gene therapy; NOV8; RNase-like; chromosome 14.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 28 /note= "This residue is Arg in a polymorphic variant"
FT Misc-difference 64 /note= "This residue is Thr in a polymorphic variant"
FT Misc-difference 69 /note= "This residue is Leu in a polymorphic variant"
FT
XX
PN WO200253742-A2.
XX
PD 11-JUL-2002.
XX
PF 07-JAN-2002; 2002WO-US000375.
XX
PR 05-JAN-2001; 2001US-0260018P.
PR 08-JAN-2001; 2001US-0260360P.
PR 28-FEB-2001; 2001US-0272411P.
PR 02-MAR-2001; 2001US-0272817P.
PR 05-JUL-2001; 2001US-0303231P.
PR 12-JUL-2001; 2001US-0305060P.
PR 10-SEP-2001; 2001US-0318405P.
PR 12-SEP-2001; 2001US-0318700P.
PR 04-JAN-2002; 2002US-00037417.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Kekuda R, Alsobrook JP, Tchernev VT, Liu X, Spytek KA;
PI Patturajan M, Grosse WM, Lepley DM, Burgess CE, Vernet CAM, Li L;
PI Gorman L, Edinger S, Sciore P, Ellerman K, Malyankar U;
PI Rothenberg M, Stone D, Boldog F, Guo X, Shenoy S, Anderson D;
PI Padigaru M, Taupier RJ, Miller CE, Eisen A;
XX
DR WPI; 2002-583619/62.
DR N-PSDB; ABQ93895.
XX

PT Novel polypeptides and nucleic acids homologous to transmembrane
PT receptor, thymosin, neuromodulin-like family of proteins for diagnosing,
PT treating cancer, atherosclerosis, neurological, skin and autoimmune
PT disorders.
XX
PS Claim 1c; Page 99; 323pp; English.
XX
CC The invention relates to 24 novel human proteins designated NOV1-NOV14
CC (ABB09501-ABB09524), collectively referred to as NOVX proteins, and
CC nucleic acids encoding them (ABQ93879-ABQ93902). NOVX proteins, and
CC nucleotides are useful in the treatment, diagnosis or prevention of NOVX-
CC associated disorders or in the manufacture of a medicament for treating
CC such disorders, with specific applications described for each of the 24
CC NOVX proteins, based on their homology to known proteins. Various
CC disorders are associated with NOVX proteins including neurological
CC disorders (e.g., Alzheimer's, Huntington's and Parkinson's diseases),
CC pain, behavioural disorders, addiction, tuberosus sclerosis, cancers
CC (e.g., colorectal cancer, leukaemia and osteosarcoma), immune disorders
CC (e.g., allergies and autoimmune diseases), myasthenia gravis, asthma,
CC various forms of arthritis, diabetes, thyroiditis, cardiovascular disease
CC (e.g., hypertension), reproductive disorders, endometriosis,
CC incontinence, psoriasis, scleroderma, alopecia, ulcers, pancreatitis,
CC cirrhosis, glomerular endothe liosis, polycystic kidney disease, endocrine
CC disorders, obesity, bacterial infections and particularly cardiomyopathy,
CC atherosclerosis, cell signal processing-related disorders and disorders
CC of metabolic pathway regulation. NOVX nucleic acids and polypeptides may
CC be used to identify cellular receptors or downstream effectors which
CC binds to a NOVX protein, and are also useful as targets for the
CC identification of small molecules that modulate or inhibit processes such
CC as neurogenesis, cell differentiation, cell motility, cellular
CC proliferation, haematopoiesis, wound healing and angiogenesis. NOVX
CC nucleic acid sequences can be used to identify a cell or tissue type and
CC are useful as a source of primers or probes for forensic biology and for
CC identifying and cloning NOVX homologues in other cell types. Cells
CC comprising NOVX nucleic acids are useful for producing non-human
CC transgenic animals which are useful for studying the function and
CC activity of NOVX proteins and for identifying and evaluating modulators
CC of NOVX activity. The present sequence represents the RNase-like protein
CC NOV8. The gene encoding NOV8 is located on chromosome 14
XX
SQ Sequence 129 AA;
Query Match 60.5%; Score 49; DB 5; Length 129;
Best Local Similarity 75.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGP S LMSW L T C 14
| | | | | | | | | |
Db 114 LTGH S LMSW L V C 125

RESULT 18
AAM29099
ID AAM29099 standard; protein; 198 AA.
XX
AC AAM29099;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #3136 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX Claim 27; SEQ ID NO 29368; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 198 AA;
Query Match 60.5%; Score 49; DB 4; Length 198;
Best Local Similarity 75.0%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLSMWLTC 14
| | | | | | | | | |
Db 183 LTGHSLSMWLVC 194
RESULT 19
AAU27672
ID AAU27672 standard; protein; 199 AA.
XX AAU27672;
AC 18-DEC-2001 (first entry)
XX Human protein APP577178.
DE Human; AFR; cytostatic; Marfan's syndrome; thrombocytopaenia; leukaemia;
XX porphyria; Gilles De La Tourette's syndrome; immunogen;
KW squamous cell carcinoma; diabetes mellitus; Grave's disease;
KW colon cancer; Alzheimer's disease; epiphyseal dysplasia.
XX Homo sapiens.
OS WO200166748-A2.
XX 13-SEP-2001.
PD 05-MAR-2001; 2001WO-US007192.
XX 03-MAR-2000; 2000US-0187221P.
PR (ZYMO) ZYMOGENETICS INC.
XX Conklin DC, Presnell SR, Adler DA;
PI WPI; 2001-589943/66.
XX N-PSDB; AAS44573.
DR Novel AFP polypeptides and polynucleotides, useful for diagnostic and
XX therapeutic purposes, in cancer therapy and for screening modulator
PT compounds.
PT Claim 2; Page 213-214; 220pp; English.
PS

XX The invention relates to novel human AFP proteins (not defined) and the
CC nucleic acids that encode them. AFP proteins are useful as standards in
CC assays of protein and protein inhibitors in both clinical and research
CC settings, as protein and amino acid supplements, including hydrolysates.
CC The nucleic acids are useful for radiation hybrid mapping. Secretary
CC fusion proteins of AFP are useful in cancer therapy, for enhancing in
CC vitro cytotoxicity, for enhancing in vivo killing of target tissues, and
CC for targeted cells or tissue inhibition or ablation. Anti-AFP antibodies
CC are useful for isolating target polypeptides by affinity purification, in
CC diagnostic assays for determining circulating or localised levels of
CC target polypeptides, for tissue typing, for cell sorting, for screening
CC expression libraries, for generating anti-idiotypic antibodies, and as
CC neutralising antibodies or as antagonists to block protein activity in
CC vitro and in vivo. AFP proteins and nucleic acids may be used to diagnose
CC or treat (e.g. by gene therapy) diseases associated with the malfunction
CC of the AFP e.g. AFP166924 and Marfan's syndrome, AFP576853/AFP39158 and
CC thrombocytopaenia, leukaemia, porphyria, Gilles De La Tourette's
CC syndrome, AFP652829 and squamous cell carcinoma, diabetes mellitus,
CC Grave's disease, AFP664311 and colon cancer, AFP253034 and Alzheimer's
CC disease and AFP686580 and epiphyseal dysplasia. Many more examples of
CC disease are given in the specification. The present sequence represents
CC an AFP of the invention
XX
SQ Sequence 199 AA;
Query Match 60.5%; Score 49; DB 4; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLSMWLTC 14
| | | | | | | | | |
Db 184 LTGHSLSMWLVC 195
RESULT 20
AAG77853
ID AAG77853 standard; protein; 199 AA.
XX AAG77853;
AC 04-JAN-2002 (first entry)
XX Human ribonuclease Zrnase1.
DE Ribonuclease; human; Zrnase1; gene therapy; tumour; inflammation; gout;
XX anaemia; Addison's disease; allergy; asthma; atherosclerosis; bronchitis;
KW Crohn's disease; ulcerative colitis; atopic dermatitis; Grave's disease;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW lupus erythematosus; multiple sclerosis; myasthenia gravis;
KW myocardial inflammation; pericardial inflammation; osteoarthritis;
KW osteoporosis; pancreatitis; rheumatoid arthritis; scleroderma;
KW Sjogren's syndrome; autoimmune thyroiditis.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Peptide 1..19
FT /label= signal_sequence
FT Protein 20..199
FT /label= mature_Zrnase1
XX WO200166710-A2.
PN 13-SEP-2001.
XX 07-MAR-2001; 2001WO-US007333.
PF 08-MAR-2000; 2000US-00522765.
XX (ZYMO) ZYMOGENETICS INC.
XX Conklin DC;
PI

XX WPI; 2001-596833/67.
DR N-PSDB; AAH78831.
XX
PT Novel isolated Zrnasei polypeptide which is a member of the human
PT ribonuclease family, useful for treating tumors, Addison's disease,
PT Crohn's disease, Grave's disease, Sjogren's syndrome and
PT glomerulonephritis.
XX
PS Claim 3; Page 2; 88pp; English.
XX
CC The sequence represents the novel human ribonuclease Zrnasei protein. The
CC invention relates to a novel isolated Zrnasei polypeptide which is a
CC member of the human ribonuclease family, having a use in gene therapy.
CC The Zrnasei polypeptide, nucleotides, antibodies and fusion proteins of
CC the invention are useful for preventing or treating a disorder associated
CC with excessive cellular proliferation such as tumour, and to prevent or
CC to treat inflammation associated disorders such as Addison's disease,
CC allergies, anaemia, asthma, atherosclerosis, bronchitis, Crohn's disease,
CC ulcerative colitis, atopic dermatitis, diabetes mellitus,
CC glomerulonephritis, gout, Grave's disease, irritable bowel syndrome,
CC lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial
CC or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune
CC thyroiditis. The polypeptide and polynucleotide may also be used as
CC educational tools in laboratory practicum kits for courses related to
CC genetics and molecular biology, protein chemistry and antibody production
CC and analysis
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 4; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSMWLVC 195

RESULT 21
AAU29085
ID AAU29085 standard; protein; 199 AA.
XX
AC AAU29085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #62.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006520.
XX
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.

PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.
PR 04-APR-2000; 2000US-0194647P.
PR 11-APR-2000; 2000US-0195975P.
PR 11-APR-2000; 2000US-0196000P.
PR 11-APR-2000; 2000US-0196187P.
PR 11-APR-2000; 2000US-0196690P.
PR 11-APR-2000; 2000US-0196820P.
PR 18-APR-2000; 2000US-0198121P.
PR 18-APR-2000; 2000US-0198585P.
PR 25-APR-2000; 2000US-0199397P.
PR 25-APR-2000; 2000US-0199550P.
PR 25-APR-2000; 2000US-0199654P.
PR 03-MAY-2000; 2000US-0201516P.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS45986.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX
PS Claim 11; Fig 124; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 4; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSMWLVC 195

RESULT 22

AAG89149
ID AAG89149 standard; protein; 199 AA.
XX
AC AAG89149;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 269.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
XX 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
XX (GEST) GENSET.
PA
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI; 2001-367870/38.
DR N-PSDB; AAH64752.
DR
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 21; Page 808; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention
XX
SQ Sequence 199 AA;
Query Match 60.5%; Score 49; DB 4; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 LVGPSLMSWLTC 14
Db 184 LTGHSLSMWLVC 195
RESULT 23
ABP43532
ID ABP43532 standard; protein; 199 AA.
XX
AC ABP43532;
XX
DT 08-AUG-2002 (first entry)

XX
DE Human secreted protein (SCEP) 56.
XX
KW Human; secreted protein; SECP; SECP expression; gene therapy;
KW protein therapy; immune system disorders; AIDS; thymic hypoplasia;
KW anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;
KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
KW developmental disorder; cell proliferative disorder; cancer.
XX
OS Homo sapiens.
XX WO200226982-A2.
XX
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030042.
XX
XX 29-SEP-2000; 2000US-0236869P.
PR 11-OCT-2000; 2000US-0239812P.
PR 12-OCT-2000; 2000US-0240108P.
PR 17-OCT-2000; 2000US-0241282P.
PR 20-OCT-2000; 2000US-0242218P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;
PI Sanjanwala MS, Walia NK, Baughn MR, Sapperstein SK, Lal P;
PI Thornton M, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;
PI Thangavelu K, Gietzen KJ, Ding L, Au-Young J, Tran B, Policky JL;
PI Lee S, Lu DAM, Burford N, Warren BA, Gururajan R, Duggan BM;
PI Honchell CD, Hafalia AJA;
XX
DR WPI; 2002-394239/42.
DR N-PSDB; ABN99415.
XX
PT New human secreted proteins, useful for diagnosing, treating or
PT preventing immune system disorders (e.g. Crohn's disease), neurological
PT disorders (e.g. Parkinson's disease), or cell proliferative disorders
PT (e.g. cancers).
XX
PS Claim 1; Page 187-188; 238pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the
CC invention are useful for treating/preventing disorders associated with
CC decreased or elevated expression of SECP. The SECP DNA and protein
CC sequences are specifically useful for treating/preventing (i.e. gene
CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic
CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders
CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);
CC developmental disorders (e.g. Down's syndrome); and cell proliferative
CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the
CC human secreted proteins (SECP) of the invention
XX
SQ Sequence 199 AA;
Query Match 60.5%; Score 49; DB 5; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 LVGPSLMSWLTC 14
Db 184 LTGHSLSMWLVC 195
RESULT 24
AAU12076
ID AAU12076 standard; protein; 199 AA.
XX
AC AAU12076;
XX
DT 23-APR-2002 (first entry)
XX

DE Human ribonuclease Zrnase1 polypeptide.
XX
KW Human; ribonuclease; Zrnase1; cellular proliferation; tumour; cancer;
KW inflammatory disorder; allergy; asthma; atherosclerosis; cytostatic;
KW autoimmune disorder; antiinflammatory; cardiant; enzyme.
XX
OS Homo sapiens.
XX
PN US2001049434-A1.
XX
PD 06-DEC-2001.
XX
PF 07-MAR-2001; 2001US-00801231.
XX
PR 08-MAR-2000; 2000US-0187917P.
XX
PA (CONK/) CONKLIN D C.
XX
PI Conklin DC;
XX
DR WPI; 2002-121402/16.
DR N-PSDB; AAS15496.
XX
PT Ribonuclease, Zrnase1, useful for preventing and treating, e.g., cancer,
PT diabetes, asthma and osteoarthritis, and to identify and to isolate
PT Zrnase1 substrates and inhibitors.
XX
PS Claim 3; Page 1; 37pp; English.
XX
CC The present invention relates to the isolation of a novel ribonuclease,
CC Zrnase1. Zrnase1 and fusion proteins comprising it can be used to prevent
CC or to treat a disorder associated with excessive cellular proliferation.
CC Examples of such disorders include various types of tumours (e.g.
CC adenocarcinoma, leukaemia, lymphoma and melanoma). In addition, Zrnase1
CC can be used to prevent or to treat inflammation associated disorders,
CC such as Addison's disease, adult respiratory distress syndrome (ARDS),
CC allergies, anaemia, asthma, atherosclerosis, bronchitis, cholecystitis,
CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
CC diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis,
CC gout, Graves' disease, hypereosinophilia, irritable bowel syndrome, lupus
CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or
CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome and
CC autoimmune thyroiditis. Zrnase1 can also be used to identify and to
CC isolate Zrnase1 substrates and inhibitors. The present sequence
CC represents human ribonuclease, Zrnase1
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 5; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTC 14
Db 184 LTGHSLSMSWLVC 195

RESULT 25
ABU58693
ID ABU58693 standard; protein; 199 AA.
XX
AC ABU58693;
XX
DT 15-APR-2003 (first entry)
XX
DE Human PRO polypeptide #294.
XX
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
OS Homo sapiens.

XX US2003027272-A1.
PN
XX 06-FEB-2003.
PD
XX 21-JUN-2002; 2002US-00176492.
PF
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.

PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.

PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTG 14
| | | | | | | | | |
Db 184 LTGHSLSWLVLC 195

RESULT 26
ABU88241
ID ABU88241 standard; protein; 199 AA.

XX AC ABU88241;
XX DT 07-JUL-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO19669.
XX KW Human; secreted and transmembrane protein: PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX OS Homo sapiens.
XX US2003032127-A1.
PN
XX
PD 13-FEB-2003.
XX

PR	10-SEP-1998;	98US-0099763P.	
PR	10-SEP-1998;	98US-0099812P.	
PR	15-SEP-1998;	98US-0100388P.	
PR	16-SEP-1998;	98US-0100662P.	
PR	16-SEP-1998;	98US-0100664P.	
PR	16-SEP-1998;	98US-0101751P.	
PR	16-SEP-1998;	98WO-US019330.	
PR	17-SEP-1998;	98US-0100683P.	
PR	17-SEP-1998;	98US-0100684P.	
PR	17-SEP-1998;	98US-0100919P.	
PR	17-SEP-1998;	98US-0100930P.	
PR	18-SEP-1998;	98US-0100849P.	
PR	18-SEP-1998;	98US-0101014P.	
PR	18-SEP-1998;	98US-0101068P.	
PR	23-SEP-1998;	98US-0101471P.	
PR	23-SEP-1998;	98US-0101472P.	
PR	23-SEP-1998;	98US-0101475P.	
PR	23-SEP-1998;	98US-0101477P.	
PR	24-SEP-1998;	98US-0101738P.	
PR	24-SEP-1998;	98US-0101739P.	
PR	24-SEP-1998;	98US-0101743P.	
PR	24-SEP-1998;	98US-0101922P.	
PR	25-SEP-1998;	98US-0101786P.	
PR	29-SEP-1998;	98US-0102207P.	
PR	29-SEP-1998;	98US-0102240P.	
PR	29-SEP-1998;	98US-0102330P.	
PR	29-SEP-1998;	98US-0102331P.	
PR	30-SEP-1998;	98US-0102487P.	
PR	30-SEP-1998;	98US-0102570P.	
PR	30-SEP-1998;	98US-0102571P.	
PR	01-OCT-1998;	98US-0102684P.	
PR	01-OCT-1998;	98US-0102687P.	
PR	02-OCT-1998;	98US-0102965P.	
PR	06-OCT-1998;	98US-0103258P.	

Query Match

Best Local Similarity

Matches

9;

Conservative

0;

Mismatches

3;

Indels

0;

Gaps

0;

QY

3

LVGPSLMSWLTC 14

| | | | | | | |

Db

184

LTGHSLMSWLVC 195

RESULT 27

ABU84556

ID

ABU84556 standard; protein; 199 AA.

XX

AC

ABU84556;

XX

DT

02-AUG-2003 (first entry)

XX

DE

Human secreted/transmembrane protein (PRO) #294.

XX

KW

Human; secreted and transmembrane protein; PRO; TNF-alpha;

KW

tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;

KW

tissue typing.

XX

OS

Homo sapiens.

XX

PN

US2003032112-A1.

XX

PD

13-FEB-2003.

XX

PF

21-JUN-2002; 2002US-00176756.

XX

PR

18-SEP-1997;

97US-0059263P.

PR

18-SEP-1997;

97US-0059266P.

PR

17-OCT-1997;

97US-0062250P.

PR

21-OCT-1997;

97US-0063486P.

PR

24-OCT-1997;

97US-0063120P.

PR

24-OCT-1997;

97US-0063121P.

PR

28-OCT-1997;

97US-0063540P.

PR	28-OCT-1997;	97US-0063541P.
PR	28-OCT-1997;	97US-0063544P.
PR	28-OCT-1997;	97US-0063564P.
PR	29-OCT-1997;	97US-0063734P.
PR	31-OCT-1997;	97US-0063870P.
PR	31-OCT-1997;	97US-0064103P.
PR	13-NOV-1997;	97US-0065311P.
PR	21-NOV-1997;	97US-0066120P.
PR	24-NOV-1997;	97US-0066466P.
PR	24-NOV-1997;	97US-0066772P.
PR	11-DEC-1997;	97US-0069335P.
PR	12-DEC-1997;	97US-0069425P.
PR	17-DEC-1997;	97US-0069870P.
PR	18-DEC-1997;	97US-0068017P.
PR	10-MAR-1998;	98US-0077450P.
PR	11-MAR-1998;	98US-0077632P.
PR	11-MAR-1998;	98US-0077649P.
PR	20-MAR-1998;	98US-0078886P.
PR	20-MAR-1998;	98US-0078939P.
PR	27-MAR-1998;	98US-0079664P.
PR	27-MAR-1998;	98US-0079786P.
PR	31-MAR-1998;	98US-0080107P.
PR	31-MAR-1998;	98US-0080194P.
PR	01-APR-1998;	98US-0080327P.
PR	01-APR-1998;	98US-0080333P.
PR	08-APR-1998;	98US-0081049P.
PR	08-APR-1998;	98US-0081070P.
PR	09-APR-1998;	98US-0081195P.
PR	15-APR-1998;	98US-0081838P.
PR	21-APR-1998;	98US-0082568P.
PR	21-APR-1998;	98US-0082569P.
PR	22-APR-1998;	98US-0082704P.
PR	22-APR-1998;	98US-0082797P.
PR	28-APR-1998;	98US-0083322P.
PR	29-APR-1998;	98US-0083495P.
PR	29-APR-1998;	98US-0083496P.
PR	29-APR-1998;	98US-0083499P.
PR	29-APR-1998;	98US-0083559P.
PR	05-MAY-1998;	98US-0084366P.
PR	06-MAY-1998;	98US-0084414P.
PR	07-MAY-1998;	98US-0084639P.
PR	07-MAY-1998;	98US-0084640P.
PR	07-MAY-1998;	98US-0084643P.
PR	15-MAY-1998;	98US-0085579P.
PR	15-MAY-1998;	98US-0085580P.
PR	15-MAY-1998;	98US-0085582P.
PR	15-MAY-1998;	98US-0085700P.
PR	18-MAY-1998;	98US-0086023P.
PR	22-MAY-1998;	98US-0086392P.
PR	28-MAY-1998;	98US-0086486P.
PR	28-MAY-1998;	98US-0087098P.
PR	28-MAY-1998;	98US-0087208P.
PR	02-JUN-1998;	98US-0087609P.
PR	03-JUN-1998;	98US-0087759P.
PR	04-JUN-1998;	98US-0088025P.
PR	04-JUN-1998;	98US-0088028P.
PR	04-JUN-1998;	98US-0088029P.
PR	04-JUN-1998;	98US-0088033P.
PR	04-JUN-1998;	98US-0088326P.
PR	05-JUN-1998;	98US-0088167P.
PR	05-JUN-1998;	98US-0088202P.
PR	05-JUN-1998;	98US-0088212P.
PR	05-JUN-1998;	98US-0088217P.
PR	09-JUN-1998;	98US-0088655P.
PR	10-JUN-1998;	98US-0088722P.
PR	10-JUN-1998;	98US-0088738P.
PR	10-JUN-1998;	98US-0088740P.
PR	10-JUN-1998;	98US-0088811P.
PR	10-JUN-1998;	98US-0088824P.
PR	10-JUN-1998;	98US-0088825P.
PR	10-JUN-1998;	98US-0088826P.
PR	11-JUN-1998;	98US-0088861P.

PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.

PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLSWLTC 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 28
ABR66430
ID ABR66430 standard; protein; 199 AA.
XX
AC ABR66430;
XX
DT 05-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulneryary; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027278-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176987.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.

PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.

PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.

PR	18-SEP-1998;	98US-0101014P.	PR	11-DEC-1997;	97US-0069333P.
PR	18-SEP-1998;	98US-0101068P.	PR	12-DEC-1997;	97US-0069425P.
PR	23-SEP-1998;	98US-0101471P.	PR	17-DEC-1997;	97US-0069870P.
PR	23-SEP-1998;	98US-0101472P.	PR	18-DEC-1997;	97US-0068017P.
PR	23-SEP-1998;	98US-0101475P.	PR	10-MAR-1998;	98US-0077450P.
PR	23-SEP-1998;	98US-0101477P.	PR	11-MAR-1998;	98US-0077632P.
PR	24-SEP-1998;	98US-0101738P.	PR	11-MAR-1998;	98US-0077649P.
PR	24-SEP-1998;	98US-0101739P.	PR	20-MAR-1998;	98US-0078886P.
PR	24-SEP-1998;	98US-0101743P.	PR	20-MAR-1998;	98US-0078939P.
PR	24-SEP-1998;	98US-0101922P.	PR	27-MAR-1998;	98US-0079664P.
PR	25-SEP-1998;	98US-0101786P.	PR	27-MAR-1998;	98US-0079786P.
PR	29-SEP-1998;	98US-0102207P.	PR	31-MAR-1998;	98US-0080107P.
PR	29-SEP-1998;	98US-0102240P.	PR	31-MAR-1998;	98US-0080194P.
PR	29-SEP-1998;	98US-0102330P.	PR	01-APR-1998;	98US-0080327P.
PR	29-SEP-1998;	98US-0102331P.	PR	01-APR-1998;	98US-0080333P.
PR	30-SEP-1998;	98US-0102487P.	PR	08-APR-1998;	98US-0081049P.
PR	30-SEP-1998;	98US-0102570P.	PR	08-APR-1998;	98US-0081070P.
PR	30-SEP-1998;	98US-0102571P.	PR	09-APR-1998;	98US-0081195P.
PR	01-OCT-1998;	98US-0102684P.	PR	15-APR-1998;	98US-0081838P.
PR	01-OCT-1998;	98US-0102687P.	PR	21-APR-1998;	98US-0082568P.
Query Match 60.5%; Score 49; DB 6; Length 199;			PR	21-APR-1998;	98US-0082569P.
Best Local Similarity 75.0%; Pred. No. 6.2;			PR	22-APR-1998;	98US-0082704P.
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			PR	22-APR-1998;	98US-0082797P.
			PR	28-APR-1998;	98US-0083322P.
QY	3 LVGPSLMSWLC 14		PR	29-APR-1998;	98US-0083495P.
			PR	29-APR-1998;	98US-0083496P.
Db	184 LTGHSLSMWLC 195		PR	29-APR-1998;	98US-0083499P.
			PR	29-APR-1998;	98US-0083559P.
RESULT 29			PR	05-MAY-1998;	98US-0084366P.
ABR65820	ABR65820 standard; protein; 199 AA.		PR	06-MAY-1998;	98US-0084414P.
XX	ABR65820;		PR	07-MAY-1998;	98US-0084639P.
AC			PR	07-MAY-1998;	98US-0084640P.
XX			PR	07-MAY-1998;	98US-0084643P.
DT	05-AUG-2003 (first entry)		PR	15-MAY-1998;	98US-0085579P.
XX			PR	15-MAY-1998;	98US-0085580P.
DE	Human secreted polypeptide PRO19669, SEQ ID NO:588.		PR	15-MAY-1998;	98US-0085582P.
XX			PR	15-MAY-1998;	98US-0085700P.
KW	Human; PRO; secreted protein; transmembrane protein;		PR	18-MAY-1998;	98US-0086023P.
KW	extracellular domain; tumour necrosis factor-alpha; TNF-alpha;		PR	22-MAY-1998;	98US-0086392P.
KW	chondrocyte; proliferation; differentiation; cartilage disorder;		PR	22-MAY-1998;	98US-0086486P.
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;		PR	28-MAY-1998;	98US-0087098P.
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;		PR	28-MAY-1998;	98US-0087208P.
KW	liver; drug screening; transgenic animal; genetic analysis;		PR	02-JUN-1998;	98US-0087609P.
KW	antiarthritic; vulnery; gene therapy.		PR	02-JUN-1998;	98US-0087759P.
XX	Homo sapiens.		PR	03-JUN-1998;	98US-0087827P.
OS			PR	04-JUN-1998;	98US-0088025P.
XX			PR	04-JUN-1998;	98US-0088028P.
PN	US2003036159-A1.		PR	04-JUN-1998;	98US-0088029P.
XX			PR	04-JUN-1998;	98US-0088033P.
PD	20-FEB-2003.		PR	04-JUN-1998;	98US-0088326P.
XX			PR	05-JUN-1998;	98US-0088167P.
PF	02-JUL-2002; 2002US-00188773.		PR	05-JUN-1998;	98US-0088202P.
XX			PR	05-JUN-1998;	98US-0088212P.
PR	18-SEP-1997;	97US-0059263P.	PR	05-JUN-1998;	98US-0088217P.
PR	18-SEP-1997;	97US-0059266P.	PR	09-JUN-1998;	98US-0088655P.
PR	17-OCT-1997;	97US-0062250P.	PR	10-JUN-1998;	98US-0088722P.
PR	21-OCT-1997;	97US-0063486P.	PR	10-JUN-1998;	98US-0088738P.
PR	24-OCT-1997;	97US-0063120P.	PR	10-JUN-1998;	98US-0088740P.
PR	24-OCT-1997;	97US-0063121P.	PR	10-JUN-1998;	98US-0088811P.
PR	28-OCT-1997;	97US-0063540P.	PR	10-JUN-1998;	98US-0088824P.
PR	28-OCT-1997;	97US-0063541P.	PR	10-JUN-1998;	98US-0088825P.
PR	28-OCT-1997;	97US-0063544P.	PR	10-JUN-1998;	98US-0088826P.
PR	28-OCT-1997;	97US-0063564P.	PR	11-JUN-1998;	98US-0088861P.
PR	29-OCT-1997;	97US-0063734P.	PR	11-JUN-1998;	98US-0088863P.
PR	31-OCT-1997;	97US-0063870P.	PR	11-JUN-1998;	98US-0088876P.
PR	31-OCT-1997;	97US-0064103P.	PR	12-JUN-1998;	98US-0089090P.
PR	13-NOV-1997;	97US-0065311P.	PR	12-JUN-1998;	98US-0089105P.
PR	21-NOV-1997;	97US-0066120P.	PR	16-JUN-1998;	98US-0089512P.
PR	24-NOV-1997;	97US-0066466P.	PR	16-JUN-1998;	98US-0089514P.
PR	24-NOV-1997;	97US-0066772P.	PR	17-JUN-1998;	98US-0089538P.
			PR	17-JUN-1998;	98US-0089598P.
			PR	17-JUN-1998;	98US-0089653P.
			PR	18-JUN-1998;	98US-0089908P.

PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.

PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 29-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTTC 14
| | | | | | | |
Db 184 LTGHSLSWLVLC 195

RESULT 30
ABU99760
ID ABU99760 standard; protein; 199 AA.
XX
AC ABU99760;

XX 09-AUG-2003 (first entry)
DT Human secreted/transmembrane protein (PRO) #294.
XX

KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.

XX Homo sapiens.

XX US2003040070-A1.

XX 27-FEB-2003.

XX 27-JUN-2002; 2002US-00184627.

XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.

PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.

PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.

PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-010062P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0103395P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLSWLTC 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 32
ABU90120
ID ABU90120 standard; protein; 199 AA.
XX
AC ABU90120;
XX
DT 11-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO19669.
XX
KW Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;
KW chondrocyte differentiation; tumour necrosis factor-alpha release;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN US2003036147-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187741.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.

PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.

PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 10-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 01-SEP-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 02-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPMSLWLTTC 14
Db 184 LTGHSLSLWLC 195

RESULT 33
ABR68369
ID ABR68369 standard; protein; 199 AA.
XX
AC ABR68369;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
chondrocyte; proliferation; differentiation; cartilage disorder;
bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
liver; drug screening; transgenic animal; genetic analysis;
antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027264-A1.
XX
PD 06-FEB-2003.
XX
PF 18-JUN-2002; 2002US-00174579.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.

PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.

PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLSWLTC 14
Db 184 LTGHSLSWLVC 195

RESULT 34
ABU96422
ID ABU96422 standard; protein; 199 AA.
XX
AC ABU96422;

XX 25-JUL-2003 (first entry)
DT Novel human secreted and transmembrane protein PRO19669.
XX
DE
XX
KW Human; secreted and transmembrane protein; PRO; transgenic animal;
KW knockout; chromosome identification; tissue typing; tumour;
KW chondrocyte proliferation; chondrocyte differentiation;
KW tumor necrosis factor-alpha release stimulator.
XX
OS Homo sapiens.
XX
PN US2003036144-A1.
XX
PD 20-FEB-2003.
XX
PF 01-JUL-2002; 2002US-00187601.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.

PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.

PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 35
ABU92853
ID ABU92853 standard; protein; 199 AA.
XX
AC ABU92853;
XX
DT 18-JUL-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; secreted protein; transmembrane protein; PRO; tumour;
proliferation; differentiation; chondrocyte cell; TNF-alpha;

KW	tumour necrosis factor-alpha; gene therapy.		
XX			
OS	Homo sapiens.		
XX			
PN	US2003036149-A1.		
XX			
PD	20-FEB-2003.		
XX			
PF			
XX	02-JUL-2002; 2002US-00187746.		
PR	18-SEP-1997;	97US-0059263P.	04-JUN-1998;
PR	18-SEP-1997;	97US-0059266P.	04-JUN-1998;
PR	17-OCT-1997;	97US-0062250P.	04-JUN-1998;
PR	21-OCT-1997;	97US-0063486P.	05-JUN-1998;
PR	24-OCT-1997;	97US-0063120P.	05-JUN-1998;
PR	24-OCT-1997;	97US-0063121P.	05-JUN-1998;
PR	28-OCT-1997;	97US-0063540P.	05-JUN-1998;
PR	28-OCT-1997;	97US-0063541P.	05-JUN-1998;
PR	28-OCT-1997;	97US-0063544P.	09-JUN-1998;
PR	28-OCT-1997;	97US-0063564P.	10-JUN-1998;
PR	29-OCT-1997;	97US-0063734P.	10-JUN-1998;
PR	31-OCT-1997;	97US-0063870P.	10-JUN-1998;
PR	31-OCT-1997;	97US-0064103P.	10-JUN-1998;
PR	13-NOV-1997;	97US-0065311P.	10-JUN-1998;
PR	21-NOV-1997;	97US-0066120P.	11-JUN-1998;
PR	24-NOV-1997;	97US-0066466P.	11-JUN-1998;
PR	24-NOV-1997;	97US-0066772P.	11-JUN-1998;
PR	11-DEC-1997;	97US-0069335P.	12-JUN-1998;
PR	12-DEC-1997;	97US-0069425P.	12-JUN-1998;
PR	17-DEC-1997;	97US-0069870P.	12-JUN-1998;
PR	18-DEC-1997;	97US-0068017P.	16-JUN-1998;
PR	10-MAR-1998;	98US-0077450P.	16-JUN-1998;
PR	11-MAR-1998;	98US-0077632P.	17-JUN-1998;
PR	11-MAR-1998;	98US-0077649P.	17-JUN-1998;
PR	20-MAR-1998;	98US-0078886P.	17-JUN-1998;
PR	20-MAR-1998;	98US-0078939P.	17-JUN-1998;
PR	27-MAR-1998;	98US-0079664P.	18-JUN-1998;
PR	27-MAR-1998;	98US-0079786P.	19-JUN-1998;
PR	31-MAR-1998;	98US-0080107P.	22-JUN-1998;
PR	31-MAR-1998;	98US-0080194P.	22-JUN-1998;
PR	01-APR-1998;	98US-0080327P.	22-JUN-1998;
PR	01-APR-1998;	98US-0080333P.	24-JUN-1998;
PR	08-APR-1998;	98US-0081049P.	24-JUN-1998;
PR	08-APR-1998;	98US-0081070P.	24-JUN-1998;
PR	09-APR-1998;	98US-0081195P.	24-JUN-1998;
PR	15-APR-1998;	98US-0081838P.	24-JUN-1998;
PR	21-APR-1998;	98US-0082568P.	25-JUN-1998;
PR	21-APR-1998;	98US-0082569P.	25-JUN-1998;
PR	22-APR-1998;	98US-0082704P.	25-JUN-1998;
PR	22-APR-1998;	98US-0082797P.	25-JUN-1998;
PR	28-APR-1998;	98US-0083322P.	25-JUN-1998;
PR	29-APR-1998;	98US-0083495P.	26-JUN-1998;
PR	29-APR-1998;	98US-0083496P.	26-JUN-1998;
PR	29-APR-1998;	98US-0083499P.	26-JUN-1998;
PR	29-APR-1998;	98US-0083559P.	26-JUN-1998;
PR	05-MAY-1998;	98US-0084366P.	01-JUL-1998;
PR	06-MAY-1998;	98US-0084414P.	01-JUL-1998;
PR	07-MAY-1998;	98US-0084639P.	01-JUL-1998;
PR	07-MAY-1998;	98US-0084640P.	02-JUL-1998;
PR	07-MAY-1998;	98US-0084643P.	02-JUL-1998;
PR	15-MAY-1998;	98US-0085579P.	02-JUL-1998;
PR	15-MAY-1998;	98US-0085580P.	02-JUL-1998;
PR	15-MAY-1998;	98US-0085582P.	02-JUL-1998;
PR	15-MAY-1998;	98US-0085700P.	24-JUL-1998;
PR	18-MAY-1998;	98US-0086023P.	04-AUG-1998;
PR	22-MAY-1998;	98US-0086392P.	10-AUG-1998;
PR	22-MAY-1998;	98US-0086486P.	10-AUG-1998;
PR	28-MAY-1998;	98US-0087098P.	17-AUG-1998;
PR	28-MAY-1998;	98US-0087208P.	17-AUG-1998;
PR	02-JUN-1998;	98US-0087609P.	17-AUG-1998;
PR	02-JUN-1998;	98US-0087759P.	18-AUG-1998;
PR	03-JUN-1998;	98US-0087827P.	18-AUG-1998;
PR	04-JUN-1998;	98US-0088025P.	18-AUG-1998;

PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTG 14
| | | | | | | | | |
Db 184 LTGHSLSMSWLC 195

RESULT 36
ABO08930
ID ABO08930 standard; protein; 199 AA.
XX
AC ABO08930;
XX
DT 17-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003044923-A1.
XX

PD 06-MAR-2003.
XX
PF 24-JUN-2002; 2002US-00179522.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.

PR	05-JUN-1998;	98US-00882112P.	PR	10-SEP-1998;	98US-0099741P.
PR	05-JUN-1998;	98US-0088217P.	PR	10-SEP-1998;	98US-0099754P.
PR	09-JUN-1998;	98US-0088655P.	PR	10-SEP-1998;	98US-0099763P.
PR	10-JUN-1998;	98US-0088722P.	PR	10-SEP-1998;	98US-0099812P.
PR	10-JUN-1998;	98US-0088738P.	PR	15-SEP-1998;	98US-0100388P.
PR	10-JUN-1998;	98US-0088740P.	PR	16-SEP-1998;	98US-0100662P.
PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088825P.	PR	16-SEP-1998;	98WO-US019330.
PR	10-JUN-1998;	98US-0088826P.	PR	17-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100684P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100919P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089090P.	PR	18-SEP-1998;	98US-0100849P.
PR	12-JUN-1998;	98US-0089105P.	PR	18-SEP-1998;	98US-0101014P.
PR	16-JUN-1998;	98US-0089512P.	PR	18-SEP-1998;	98US-0101068P.
PR	16-JUN-1998;	98US-0089514P.	PR	23-SEP-1998;	98US-0101471P.
PR	17-JUN-1998;	98US-0089538P.	PR	23-SEP-1998;	98US-0101472P.
PR	17-JUN-1998;	98US-0089598P.	PR	23-SEP-1998;	98US-0101475P.
PR	17-JUN-1998;	98US-0089653P.	PR	23-SEP-1998;	98US-0101477P.
PR	18-JUN-1998;	98US-0089908P.	PR	24-SEP-1998;	98US-0101738P.
PR	19-JUN-1998;	98US-0089952P.	PR	24-SEP-1998;	98US-0101739P.
PR	22-JUN-1998;	98US-0090246P.	PR	24-SEP-1998;	98US-0101743P.
PR	22-JUN-1998;	98US-0090252P.	PR	24-SEP-1998;	98US-0101922P.
PR	22-JUN-1998;	98US-0090254P.	PR	25-SEP-1998;	98US-0101786P.
PR	24-JUN-1998;	98US-0090429P.	PR	29-SEP-1998;	98US-0102207P.
PR	24-JUN-1998;	98US-0090435P.	PR	29-SEP-1998;	98US-0102240P.
PR	24-JUN-1998;	98US-0090444P.	PR	29-SEP-1998;	98US-0102330P.
PR	24-JUN-1998;	98US-0090461P.	PR	29-SEP-1998;	98US-0102331P.
PR	24-JUN-1998;	98US-0090535P.	PR	30-SEP-1998;	98US-0102487P.
PR	24-JUN-1998;	98US-0090540P.	PR	30-SEP-1998;	98US-0102570P.
PR	25-JUN-1998;	98US-0090676P.	PR	30-SEP-1998;	98US-0102571P.
PR	25-JUN-1998;	98US-0090678P.	PR	01-OCT-1998;	98US-0102684P.
PR	25-JUN-1998;	98US-0090688P.	PR	01-OCT-1998;	98US-0102687P.
PR	25-JUN-1998;	98US-0090690P.	PR	02-OCT-1998;	98US-0102965P.
PR	25-JUN-1998;	98US-0090694P.	PR	06-OCT-1998;	98US-0103258P.
PR	25-JUN-1998;	98US-0090695P.	PR	06-OCT-1998;	98US-0103449P.
PR	25-JUN-1998;	98US-0090696P.	Query Match 60.5%; Score 49; DB 6; Length 199;		
PR	26-JUN-1998;	98US-00105413.	Best Local Similarity 75.0%; Pred. No. 6.2;		
PR	26-JUN-1998;	98US-0090862P.	Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
PR	26-JUN-1998;	98US-0090863P.			
PR	26-JUN-1998;	98US-0091010P.			
PR	01-JUL-1998;	98US-0091359P.	QY	3	LVGPSLMSWLT
PR	01-JUL-1998;	98US-0091544P.	Db	184	LTGHSLSWLVC
PR	02-JUL-1998;	98US-0091478P.			
PR	02-JUL-1998;	98US-0091486P.			
PR	02-JUL-1998;	98US-0091626P.			
PR	02-JUL-1998;	98US-0091628P.			
PR	02-JUL-1998;	98US-0091632P.			
PR	24-JUL-1998;	98US-0094006P.			
PR	04-AUG-1998;	98US-0095282P.			
PR	10-AUG-1998;	98US-0095998P.			
PR	10-AUG-1998;	98US-0096012P.			
PR	17-AUG-1998;	98US-0096757P.			
PR	17-AUG-1998;	98US-0096766P.			
PR	17-AUG-1998;	98US-0096867P.			
PR	17-AUG-1998;	98US-0096891P.			
PR	17-AUG-1998;	98US-0096897P.			
PR	18-AUG-1998;	98US-0096949P.			
PR	18-AUG-1998;	98US-0096959P.			
PR	18-AUG-1998;	98US-0097022P.			
PR	26-AUG-1998;	98US-0097952P.			
PR	26-AUG-1998;	98US-0097954P.			
PR	26-AUG-1998;	98US-0097955P.			
PR	26-AUG-1998;	98US-0097971P.			
PR	26-AUG-1998;	98US-0097974P.			
PR	26-AUG-1998;	98US-0098014P.			
PR	01-SEP-1998;	98US-0098716P.			
PR	01-SEP-1998;	98US-0098723P.			
PR	02-SEP-1998;	98US-0098803P.			
PR	02-SEP-1998;	98US-0098821P.			
PR	02-SEP-1998;	98US-0098843P.			
PR	09-SEP-1998;	98US-0099602P.			
RESULT 37					
ABO02982					
ID	ABO02982 standard; protein; 199 AA.				
XX					
AC	ABO02982;				
XX					
DT	09-AUG-2003 (first entry)				
XX					
DE	Human secreted/transmembrane protein (PRO) #294.				
XX					
KW	Human; secreted and transmembrane protein; PRO; TNF-alpha;				
KW	tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;				
KW	tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;				
KW	prostate tumour; rectal tumour; cervical tumour; liver tumour.				
XX					
OS	Homo sapiens.				
XX					
PN	US2003040062-A1.				
XX					
PD	27-FEB-2003.				
XX					
PF	25-JUN-2002; 2002US-00180545.				
XX					
PR	18-SEP-1997; 97US-0059263P.				
PR	18-SEP-1997; 97US-0059266P.				
PR	17-OCT-1997; 97US-0062250P.				

PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069333P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.

PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 30-SEP-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 01-OCT-1998; 98US-0102965P.
PR 02-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSMWLVC 195

RESULT 38
ABR75136
ID ABR75136 standard; protein; 199 AA.

XX AC ABR75136;

XX DT 10-SEP-2003 (first entry)

XX DE Human secreted polypeptide PRO19669, SEQ ID NO:588.

XX KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnerary; gene therapy.

XX OS Homo sapiens.

XX PN US2003040056-A1.

XX PD 27-FEB-2003.

XX PF 21-JUN-2002; 2002US-00176916.

XX PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.

PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-007886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.

PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.

PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLSWLTC 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 39
ABR94898
ID ABR94898 standard; protein; 199 AA.
XX
AC ABR94898;
XX
DT 13-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX

KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.

XX Homo sapiens.

OS US2003044926-A1.

XX
PN 06-MAR-2003.

PD 26-JUN-2002; 2002US-00183015.

XX
PF 18-SEP-1997; 97US-0059263P.
XX 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.

PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069333P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087609P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.

PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.

PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTG 14
| | | | | | | |
Db 184 LTGHSLSWLVG 195

RESULT 40
ABU85871
ID ABU85871 standard; protein; 199 AA.
XX
AC ABU85871;
XX
DT 02-JUL-2003 (first entry)
XX
DE Human PRO polypeptide #294.
XX
KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor alpha; TNF-alpha; chondrocyte cell; tumour;
KW cytostatic.
XX
OS Homo sapiens.
XX
OS
PN US2003036140-A1.
XX
PD 20-FEB-2003.
XX
PF 01-JUL-2002; 2002US-00187588.
XX

PR 26-JUN-1998; 98US-00105413.
PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.
PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.

(GETH) GENENTECH INC.
Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
WPI; 2003-332028/31.
N-PSDB; ACA73064.
Three hundred and five nucleic acids encoding PRO polypeptides, useful
for the manufacture of a medicament for diagnosing or treating tumor.
Claim 11; Fig 588; 707pp; English.
The invention relates to human PRO polypeptides (secreted and
transmembrane polypeptides) and the PRO polynucleotides encoding them.
The invention also relates to a method for stimulating the release of
tumour necrosis factor alpha (TNF-alpha) from human blood by contacting
the blood with a sequence of the invention, a method for stimulating the
proliferation or differentiation of chondrocyte cells by contacting the
cells with a PRO polypeptide and a method for detecting the presence of a
tumour in a mammal. The polypeptides and polynucleotides are useful for
the manufacture of a medicament for diagnosing or treating a tumour in a
mammal. Sequences ABU85578-ABU85882 represent human PRO polypeptides of
the invention. Note: The sequence data for this patent is also available
in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 199 AA;

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLC 14
Db 184 LTGHSLSWLC 195

RESULT 42
ABU98246

ID ABU98246 standard; protein; 199 AA.
XX
AC ABU98246;
XX
DT 30-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO19669.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome identification.
XX
OS Homo sapiens.
XX
PN US2003017544-A1.
XX
PD 23-JAN-2003.
XX
PF 21-JUN-2002; 2002US-00176915.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.

PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089553P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.

PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLT 14
Db 184 LTGHSLMSWLC 195

RESULT 43
ABU91952
ID ABU91952 standard; protein; 199 AA.
XX AC ABU91952;
XX 11-AUG-2003 (first entry)
DT Novel human secreted and transmembrane protein PRO19669.
XX DE

XX Human; gene therapy; chromosome identification; tissue typing.

OS Homo sapiens.

PN US2003027277-A1.

XX PD 06-FEB-2003.

XX PF 21-JUN-2002; 2002US-00176985.

PR 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062250P.

PR 21-OCT-1997; 97US-0063486P.

PR 24-OCT-1997; 97US-0063120P.

PR 24-OCT-1997; 97US-0063121P.

PR 28-OCT-1997; 97US-0063540P.

PR 28-OCT-1997; 97US-0063541P.

PR 28-OCT-1997; 97US-0063544P.

PR 28-OCT-1997; 97US-0063564P.

PR 29-OCT-1997; 97US-0063734P.

PR 31-OCT-1997; 97US-0063870P.

PR 31-OCT-1997; 97US-0064103P.

PR 13-NOV-1997; 97US-0065311P.

PR 21-NOV-1997; 97US-0066120P.

PR 24-NOV-1997; 97US-0066466P.

PR 24-NOV-1997; 97US-0066772P.

PR 11-DEC-1997; 97US-0069335P.

PR 12-DEC-1997; 97US-0069425P.

PR 17-DEC-1997; 97US-0069870P.

PR 18-DEC-1997; 97US-0068017P.

PR 10-MAR-1998; 98US-0077450P.

PR 11-MAR-1998; 98US-0077632P.

PR 11-MAR-1998; 98US-0077649P.

PR 20-MAR-1998; 98US-0078886P.

PR 20-MAR-1998; 98US-0078939P.

PR 27-MAR-1998; 98US-0079664P.

PR 27-MAR-1998; 98US-0079786P.

PR 31-MAR-1998; 98US-0080107P.

PR 31-MAR-1998; 98US-0080194P.

PR 01-APR-1998; 98US-0080327P.

PR 01-APR-1998; 98US-0080333P.

PR 08-APR-1998; 98US-0081049P.

PR 09-APR-1998; 98US-0081070P.

PR 04-JUN-1998; 98US-0088025P.

PR 04-JUN-1998; 98US-0088028P.

PR 04-JUN-1998; 98US-0088029P.

PR 04-JUN-1998; 98US-0088033P.

PR 04-JUN-1998; 98US-0088326P.

PR 05-JUN-1998; 98US-0088167P.

PR 05-JUN-1998; 98US-0088202P.

PR 05-JUN-1998; 98US-0088212P.

PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.

PR 10-JUN-1998; 98US-0088722P.

PR 10-JUN-1998; 98US-0088738P.

PR 10-JUN-1998; 98US-0088740P.

PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.

PR 10-JUN-1998; 98US-0088825P.

PR 10-JUN-1998; 98US-0088826P.

PR 11-JUN-1998; 98US-0088861P.

PR 11-JUN-1998; 98US-0088863P.

PR 11-JUN-1998; 98US-0088876P.

PR 12-JUN-1998; 98US-0089090P.

PR 12-JUN-1998; 98US-0089105P.

PR 16-JUN-1998; 98US-0089512P.

PR 16-JUN-1998; 98US-0089514P.

PR 17-JUN-1998; 98US-0089538P.

PR 17-JUN-1998; 98US-0089598P.

PR 17-JUN-1998; 98US-0089653P.

PR 18-JUN-1998; 98US-0089908P.

PR 19-JUN-1998; 98US-0089952P.

PR 22-JUN-1998; 98US-0090246P.

PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.

PR 24-JUN-1998; 98US-0090429P.

PR 24-JUN-1998; 98US-0090435P.

PR 24-JUN-1998; 98US-0090444P.

PR 24-JUN-1998; 98US-0090461P.

PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090540P.

PR 25-JUN-1998; 98US-0090676P.

PR 25-JUN-1998; 98US-0090678P.

PR 25-JUN-1998; 98US-0090688P.

PR 25-JUN-1998; 98US-0090690P.

PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.

PR 25-JUN-1998; 98US-0090696P.

PR 26-JUN-1998; 98US-00105413.

PR 26-JUN-1998; 98US-0090862P.

PR 26-JUN-1998; 98US-0090863P.

PR 26-JUN-1998; 98US-0091010P.

PR 01-JUL-1998; 98US-0091359P.

PR 01-JUL-1998; 98US-0091544P.

PR 02-JUL-1998; 98US-0091478P.

PR 02-JUL-1998; 98US-0091486P.

PR 02-JUL-1998; 98US-0091626P.

PR 02-JUL-1998; 98US-0091628P.

PR 02-JUL-1998; 98US-0091632P.

PR 24-JUL-1998; 98US-0094006P.

PR 04-AUG-1998; 98US-0095282P.

PR 10-AUG-1998; 98US-0095998P.

PR 10-AUG-1998; 98US-0096012P.

PR 17-AUG-1998; 98US-0096757P.

PR 17-AUG-1998; 98US-0096766P.

PR 17-AUG-1998; 98US-0096867P.

PR 17-AUG-1998; 98US-0096891P.

PR 17-AUG-1998; 98US-0096897P.

PR 18-AUG-1998; 98US-0096949P.

PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.

PR 26-AUG-1998; 98US-0097952P.

PR 26-AUG-1998; 98US-0097954P.

PR 26-AUG-1998; 98US-0097955P.

PR 26-AUG-1998; 98US-0097971P.

PR 26-AUG-1998; 98US-0097974P.

PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101058P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 45
ABU86486
ID ABU86486 standard; protein; 199 AA.

XX ABU86486;
XX
DT 01-JUL-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; immunogen; secreted protein; transmembrane protein; PRO; tumour;
KW proliferation; differentiation; chondrocyte cells;
KW tumour necrosis factor-alpha; TNF-alpha; blood; gene therapy.

OS Homo sapiens.
XX
PN US2003036146-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-00187603.
XX
PR 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.
PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.
XX
PA (GETH) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
PI
XX WPI; 2003-332034/31.
DR N-PSDB; ACA73678.
DR
XX

PT Three hundred and five nucleic acids encoding PRO polypeptides, useful in
PT gene therapy, chromosome identification, tissue typing, and for detecting
PT the presence of tumor in a mammal.
XX
PS Claim 11; Fig 588; 707pp; English.
XX
CC The invention relates to three hundred and five nucleic acids encoding
CC PRO polypeptides (secreted and transmembrane), sequences 80% identical to
CC them, or encoding a PRO polypeptide lacking its associated signal peptide
CC or an extracellular domain of the PRO polypeptide, with or lacking its
CC associated signal peptide. Also included are the encoded PRO proteins,
CC PRO expression vectors, host cells transformed with the vector (used to
CC produce PRO proteins), a chimaeric molecule comprising the PRO
CC polypeptide fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, a method for stimulating the release of tumor necrosis factor
CC alpha (TNF-alpha) from human blood (by contacting the blood with PRO1079,
CC PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or
CC PRO4333), a method for stimulating the proliferation or differentiation
CC of chondrocyte cells by contacting the cells with a PRO6029 polypeptide,
CC a method for detecting the presence of tumour in a mammal and an
CC oligonucleotide probe derived from any of the nucleotide sequences cited
CC above. The PRO polypeptide or anti-PRO antibody is useful for preparing a
CC medicament for treating a condition that is responsive to the PRO
CC polypeptide or anti-PRO antibody. The PRO nucleotide sequences are useful
CC as hybridisation probes in chromosome and gene mapping, or in generating
CC antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO
CC polypeptides, in assays to identify other proteins or molecules involved
CC in a binding reaction, to generate transgenic animals or knockout
CC animals, which in turn are useful in the development and screening of
CC therapeutically useful reagents, for chromosome identification, and
CC tissue typing. The PRO polypeptides and nucleic acid molecules are also
CC useful for detecting the presence of a tumour in a mammal, stimulating the
CC proliferation or differentiation of chondrocyte cells, stimulating the
CC release of tumour necrosis factor-alpha from human blood, in gene
CC therapy, or as molecular weight markers for protein electrophoresis
CC purposes. The anti-PRO antibodies may be used in diagnostic assays for
CC PRO, or for the affinity purification of PRO from recombinant cell
CC culture or natural sources. The present sequence represents a PRO protein
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLSMWLTC 14
Db 184 LTGHSLMSWLVC 195

RESULT 46
ABU67699
ID ABU67699 standard; protein; 199 AA.
XX
AC ABU67699;
XX 29-MAY-2003 (first entry)
DT
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX
OS Homo sapiens.
XX
PN US2003036162-A1.
XX
PD 20-FEB-2003.
XX
PF 12-JUL-2002; 2002US-00194423.
XX
PR 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.
PR 06-NOV-1998; 98US-00187368.
PR 01-DEC-1998; 98WO-US025108.
PR 07-DEC-1998; 98US-00202054.
PR 03-MAR-1999; 99US-00254311.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 25-MAY-2001; 2001US-00854280.
PR 01-JUN-2001; 2001US-00866028.
PR 05-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.

(GETH) GENENTECH INC.

XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-332039/31.
DR N-PSDB; ACA05993.
XX

PT New secreted and transmembrane PRO polypeptides and nucleic acids, useful
PT in gene therapy, in chromosome and gene mapping, as chromosome markers,
PT in tissue typing, and in chromosome identification.

XX Claim 11; Fig 588; 706pp; English.

CC The invention discloses human nucleic acids encoding secreted and
CC transmembrane (PRO) polypeptides. Also disclosed is an antibody that
CC specifically binds to the PRO polypeptide, a method for stimulating the
CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by
CC contacting the blood a PRO polypeptide, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells by contacting the
CC cells with a PRO polypeptide, a method for detecting the presence of a
CC tumour in a mammal and an oligonucleotide probe derived from any of the
CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
CC preparing PRO polypeptides by recombinant techniques and in gene therapy
CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
CC as molecular weight markers for protein electrophoresis purposes, for
CC chromosome identification, as chromosome markers, as therapeutic agents,
CC for stimulating the release of TNF-alpha from human blood, for
CC stimulating the proliferation or differentiation of chondrocytes and
CC detecting the presence of a tumour. The PRO polypeptides and nucleic
CC acids may also be used diagnostically for tissue typing. The sequences
CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 6; Length 199;

Best Local Similarity 75.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLSMSWLVC 195

RESULT 47

ABU80727

ID ABU80727 standard; protein; 199 AA.

XX AC ABU80727;

XX DT 23-JUN-2003 (first entry)

XX DE Human PRO protein #294.

XX KW Human; tumour; adrenal; lung; colon; breast; prostate; rectal; cervical;
KW liver; PRO; gene therapy.

XX OS Homo sapiens.

XX PN US2003036137-A1.

XX PD 20-FEB-2003.

XX PF 27-JUN-2002; 2002US-00184640.

XX PR 26-JUN-1998; 98US-00105413.

XX PR 16-SEP-1998; 98WO-US019330.

XX PR 07-OCT-1998; 98US-00168978.

XX PR 07-OCT-1998; 98WO-US021141.

XX PR 06-NOV-1998; 98US-00187368.

XX PR 01-DEC-1998; 98WO-US025108.

XX PR 07-DEC-1998; 98US-00202054.

XX PR 03-MAR-1999; 99US-00254311.

XX PR 08-MAR-1999; 99WO-US005028.

XX PR 14-MAY-1999; 99US-00311832.

XX PR 14-MAY-1999; 99WO-US010733.

XX PR 02-JUN-1999; 99WO-US012252.

XX PR 25-AUG-1999; 99US-00380137.

XX PR 25-AUG-1999; 99US-00380138.

XX PR 25-AUG-1999; 99US-00380139.

PR 25-AUG-1999; 99US-00380142.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028551.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 29-AUG-2001; 2001WO-US027099.
PR 04-SEP-2001; 2001US-00946374.
PR 15-JAN-2002; 2002US-00052586.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-342038/32.
N-PSDB; ACA66827.

Three hundred and five nucleic acids encoding secreted and transmembrane
PRO polypeptides, useful for the diagnosis, prevention and/or treatment
of tumors, such as adrenal, lung, colon, breast, prostate, rectal,
cervical or liver tumors.

Claim 11; Fig 588; 708pp; English.

The invention relates to three hundred and five nucleic acids encoding
PRO polypeptides (secreted and transmembrane). Methods and compositions
of the present invention are useful for the diagnosis, prevention and/or
treatment of tumors, such as adrenal, lung, colon, breast, prostate,
rectal, cervical or liver tumours. The PRO polypeptides are also useful
as molecular weight markers, or for chromosome identification. The PRO
genes are useful as hybridisation probes, or for screening libraries of

CC human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene
CC therapy, particularly for replacing a defective gene. The present
CC sequence represents a human PRO polypeptide of the invention
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSILMSWLTC 14
Db 184 LTGHSILMSWLVC 195

RESULT 48
ABR99645
ID ABR99645 standard; protein; 199 AA.
XX
AC ABR99645;
XX
DT 18-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003040063-A1.
XX
PD 27-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183006.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.

PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 8WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match

60.5%; Score 49; DB 6; Length 199;

Best Local Similarity 75.0%; Pred. No. 6.2;										
Matches		9;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
QY	3 LVGPSLSWSLTC 14									
Db	184 LTGHSLSWSLVC 195									
RESULT 49										
ABR99035										
ID	ABR99035 standard; protein; 199 AA.									
XX	XX									
AC	ABR99035;									
XX	XX									
DT	17-SEP-2003 (first entry)									
DE	Human secreted polypeptide PRO19669, SEQ ID NO:588.									
XX	XX									
KW	Human; PRO; secreted protein; transmembrane protein; TNF-alpha;									
KW	extracellular domain; tumour necrosis factor-alpha; cartilage disorder;									
KW	chondrocyte; proliferation; differentiation; cancer; tumour; diagnosis;									
KW	bone disorder; arthritis; sports injury; cancer; kidney; rectum; cervix;									
KW	adrenal tumour; lung; colon; breast; prostate; genetic analysis;									
KW	liver; drug screening; transgenic animal; gene therapy.									
KW	antiarthritic; vulnerary; gene therapy.									
XX	XX									
OS	Homo sapiens.									
XX	XX									
PN	US2003040064-A1.									
XX	XX									
PD	27-FEB-2003.									
XX	XX									
PF	26-JUN-2002; 2002US-00183008.									
XX	XX									
PR	18-SEP-1997; 97US-0059263P.									
PR	18-SEP-1997; 97US-0059266P.									
PR	17-OCT-1997; 97US-0062250P.									
PR	21-OCT-1997; 97US-0063486P.									
PR	24-OCT-1997; 97US-0063120P.									
PR	24-OCT-1997; 97US-0063121P.									
PR	28-OCT-1997; 97US-0063540P.									
PR	28-OCT-1997; 97US-0063541P.									
PR	28-OCT-1997; 97US-0063544P.									
PR	28-OCT-1997; 97US-0063564P.									
PR	29-OCT-1997; 97US-0063734P.									
PR	31-OCT-1997; 97US-0063870P.									
PR	31-OCT-1997; 97US-0064103P.									
PR	13-NOV-1997; 97US-0055311P.									
PR	21-NOV-1997; 97US-0066120P.									
PR	24-NOV-1997; 97US-0066466P.									
PR	24-NOV-1997; 97US-0066772P.									
PR	11-DEC-1997; 97US-0069335P.									
PR	12-DEC-1997; 97US-0069425P.									
PR	17-DEC-1997; 97US-0069870P.									
PR	18-DEC-1997; 97US-0068017P.									
PR	10-MAR-1998; 98US-0077450P.									
PR	11-MAR-1998; 98US-0077632P.									
PR	11-MAR-1998; 98US-0077649P.									
PR	20-MAR-1998; 98US-0078886P.									
PR	20-MAR-1998; 98US-0078939P.									
PR	27-MAR-1998; 98US-0079664P.									
PR	27-MAR-1998; 98US-0079786P.									
PR	31-MAR-1998; 98US-0080107P.									
PR	31-MAR-1998; 98US-0080194P.									
PR	01-APR-1998; 98US-0080327P.									
PR	01-APR-1998; 98US-0080333P.									
PR	08-APR-1998; 98US-0081049P.									
PR	08-APR-1998; 98US-0081070P.									
PR	09-APR-1998; 98US-0081195P.									
PR	15-APR-1998; 98US-0081838P.									
PR	21-APR-1998; 98US-0082568P.									
PR	21-APR-1998; 98US-0082569P.									
PR	22-APR-1998; 98US-0082704P.									
PR	22-APR-1998; 98US-0082797P.									
PR	28-APR-1998; 98US-0083322P.									
PR	29-APR-1998; 98US-0083495P.									
PR	29-APR-1998; 98US-0083496P.									
PR	29-APR-1998; 98US-0083499P.									
PR	29-APR-1998; 98US-0083559P.									
PR	05-MAY-1998; 98US-0084366P.									
PR	06-MAY-1998; 98US-0084414P.									
PR	07-MAY-1998; 98US-0084639P.									
PR	07-MAY-1998; 98US-0084640P.									
PR	07-MAY-1998; 98US-0084643P.									
PR	15-MAY-1998; 98US-0085579P.									
PR	15-MAY-1998; 98US-0085580P.									
PR	15-MAY-1998; 98US-0085582P.									
PR	15-MAY-1998; 98US-0085700P.									
PR	18-MAY-1998; 98US-0086023P.									
PR	22-MAY-1998; 98US-0086392P.									
PR	22-MAY-1998; 98US-0086486P.									
PR	28-MAY-1998; 98US-0087098P.									
PR	28-MAY-1998; 98US-0087208P.									
PR	02-JUN-1998; 98US-0087609P.									
PR	02-JUN-1998; 98US-0087759P.									
PR	03-JUN-1998; 98US-0087827P.									
PR	04-JUN-1998; 98US-0088025P.									
PR	04-JUN-1998; 98US-0088028P.									
PR	04-JUN-1998; 98US-0088029P.									
PR	04-JUN-1998; 98US-0088033P.									
PR	04-JUN-1998; 98US-0088326P.									
PR	05-JUN-1998; 98US-0088167P.									
PR	05-JUN-1998; 98US-0088202P.									
PR	05-JUN-1998; 98US-0088212P.									
PR	05-JUN-1998; 98US-0088217P.									
PR	09-JUN-1998; 98US-0088655P.									
PR	10-JUN-1998; 98US-0088722P.									
PR	10-JUN-1998; 98US-0088738P.									
PR	10-JUN-1998; 98US-0088740P.									
PR	10-JUN-1998; 98US-0088811P.									
PR	10-JUN-1998; 98US-0088824P.									
PR	10-JUN-1998; 98US-0088825P.									
PR	10-JUN-1998; 98US-0088826P.									
PR	11-JUN-1998; 98US-0088861P.									
PR	11-JUN-1998; 98US-0088863P.									
PR	11-JUN-1998; 98US-0088876P.									
PR	12-JUN-1998; 98US-0089090P.									
PR	12-JUN-1998; 98US-0089105P.									
PR	16-JUN-1998; 98US-0089512P.									
PR	16-JUN-1998; 98US-0089514P.									
PR	17-JUN-1998; 98US-0089538P.									
PR	17-JUN-1998; 98US-0089598P.									
PR	17-JUN-1998; 98US-0089653P.									
PR	18-JUN-1998; 98US-0089908P.									
PR	19-JUN-1998; 98US-0089952P.									
PR	22-JUN-1998; 98US-0090246P.									
PR	22-JUN-1998; 98US-0090252P.									
PR	22-JUN-1998; 98US-0090254P.									
PR	24-JUN-1998; 98US-0090429P.									
PR	24-JUN-1998; 98US-0090435P.									
PR	24-JUN-1998; 98US-0090444P.									
PR	24-JUN-1998; 98US-0090461P.									
PR	24-JUN-1998; 98US-0090535P.									
PR	24-JUN-1998; 98US-0090540P.									
PR	25-JUN-1998; 98US-0090676P.									
PR	25-JUN-1998; 98US-0090678P.									
PR	25-JUN-1998; 98US-0090688P.									
PR	25-JUN-1998; 98US-0090690P.									
PR	25-JUN-1998; 98US-0090694P.									
PR	25-JUN-1998; 98US-0090695P.									
PR	25-JUN-1998; 98US-0090696P.									
PR	26-JUN-1998; 98US-00105413.									
PR	26-JUN-1998; 98US-0090862P.									
PR	26-JUN-1998; 98US-0090863P.									
PR	26-JUN-1998; 98US-0091010P.									
PR	01-JUL-1998; 98US-0091359P.									

PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLSWLTLC 14
Db 184 LTGHSLMSWLC 195

RESULT 50
ABO16558
ID ABO16558 standard; protein; 199 AA.
XX
AC ABO16558;
XX
DT 25-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003027267-A1.
XX
PD 06-FEB-2003.
XX
PF 19-JUN-2002; 2002US-00175739.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.

PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.

PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTLC 14
| | | | | | | | | |
Db 184 LTGHSLMSWLVC 195

Search completed: May 12, 2006, 10:36:06
Job time : 130.809 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16.9915 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-67
Perfect score: 81
Sequence: 1 CELVGPSLMSWLTC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	56.8	365	T30343	irp3 protein - Yer
2	46	56.8	365	T17441	hypothetical prote
3	46	56.8	366	AI0232	yersiniabactin bio
4	45	55.6	907	AB1885	hypothetical prote
5	44	54.3	1132	A35098	MHC class III hist
6	43	53.1	224	T35918	hypothetical prote
7	42.5	52.5	261	C55581	nasD protein - Kle
8	42	51.9	538	T00424	probable Na+/Ca2+
9	42	51.9	686	S33475	hypothetical prote
10	42	51.9	786	G88065	protein T16A1.2 [i
11	42	51.9	869	S49844	probable membrane
12	41	50.6	263	EQBOA	enkephalin e-12 pr
13	41	50.6	290	A87380	conserved hypothet
14	41	50.6	361	H71631	queuine tRNA-ribos
15	41	50.6	389	S26414	naringenin-chalcon
16	41	50.6	703	T03888	Na+/Ca2+,K+-exchan
17	41	50.6	760	T34414	hypothetical prote
18	40	49.4	84	AB2301	hypothetical prote
19	40	49.4	372	T02735	hypothetical prote
20	40	49.4	410	DEPSXA	3-methyl-2-oxobuta
21	40	49.4	410	C83365	2-oxoisovalerate d
22	40	49.4	459	T31134	naphthalene dioxyg
23	40	49.4	651	T03889	Na+/Ca2+,K+-exchan
24	40	49.4	1068	T04112	pol protein homolo
25	39	48.1	266	S54440	hemin-specific ATP
26	39	48.1	306	C85224	probable LTR retro
27	39	48.1	306	T04759	hypothetical prote
28	39	48.1	465	T48374	UDPg glucosyltrans
29	39	48.1	735	S46830	urea transport pro

30	46.9	112	2	S75152	hypothetical prote
31	46.9	219	2	A46097	GPI-anchor biosynt
32	46.9	247	2	C90848	hypothetical prote
33	46.9	247	2	B85706	hypothetical prote
34	46.9	247	2	B64873	probable membrane
35	46.9	267	1	EQHUA	enkephalin precurs
36	46.9	286	2	C95847	hypothetical trans
37	46.9	342	2	T09355	hypothetical prote
38	46.9	357	2	T12379	NADH2 dehydrogenas
39	46.9	437	1	B29336	ubiquinol-cytochro
40	46.9	441	2	B69004	hypothetical prote
41	46.9	455	2	AG0308	isochorismate synt
42	46.9	472	2	AI3367	multidrug resistanc
43	46.9	493	2	T48630	high affinity nitro
44	46.9	499	2	H83254	probable MFS trans
45	46.9	541	1	A70022	multidrug-efflux t
46	46.9	594	2	T22072	hypothetical prote
47	46.9	644	2	T22067	hypothetical prote
48	46.9	718	2	H75485	hypothetical prote
49	46.9	846	2	T04970	probable potassium
50	46.9	941	2	B96553	hypothetical prote
51	46.9	1149	2	I38006	M130 antigen precu
52	46.9	1151	2	I38004	M130 antigen precu
53	46.9	1156	2	I38005	M130 antigen precu
54	37.5	571	2	A75499	DNA primase - Dein
55	45.7	248	2	H83111	hypothetical prote
56	45.7	260	2	S72748	B1177_F3_136 prote
57	45.7	266	2	T12073	ATP-binding protei
58	45.7	266	2	AC0035	hemin transport sy
59	45.7	297	2	AC0615	conserved hypothet
60	45.7	297	2	B90754	hypothetical prote
61	45.7	297	2	F64831	hypothetical prote
62	45.7	297	2	H85617	hypothetical prote
63	45.7	333	2	D95968	probable amino aci
64	45.7	371	2	D96501	hypothetical prote
65	45.7	378	2	D83381	hypothetical prote
66	45.7	387	2	T40309	cytochrome c oxida
67	45.7	445	2	T01591	hypothetical prote
68	45.7	602	2	T48697	hypothetical prote
69	45.7	868	2	T25716	hypothetical prote
70	45.7	1121	2	T25715	hypothetical prote
71	45.7	1240	2	T04833	hypothetical prote
72	45.7	1309	2	H96650	protein T3P18.3 [i
73	45.7	1415	2	C86438	hypothetical prote
74	45.7	1796	2	AC1895	serine/threonine k
75	45.7	3856	2	T51174	ataxia-telangiecta
76	36.5	45.1	229	H72110	ribulose-phosphate
77	36.5	45.1	229	A86514	ribulose-P epimera
78	36.5	45.1	261	A60091	per region circadi
79	36.5	45.1	1572	T00027	brain-specific ang
80	36	44.4	166	F96786	protein F10A5.11 [
81	36	44.4	199	T45543	hypothetical prote
82	36	44.4	204	T28939	hypothetical prote
83	36	44.4	233	S74430	hypothetical prote
84	36	44.4	247	T45502	hypothetical prote
85	36	44.4	304	AD1029	probable membrane
86	36	44.4	313	S75303	hypothetical prote
87	36	44.4	315	T11388	NADH2 dehydrogenas
88	36	44.4	330	S32773	alpha-antigen - My
89	36	44.4	335	T04029	hypothetical prote
90	36	44.4	370	B35255	chloromuconate cyc
91	36	44.4	382	T11219	ubiquinol-cytochro
92	36	44.4	446	D81058	hypothetical prote
93	36	44.4	455	H69752	probable glucarat
94	36	44.4	477	B75409	multidrug-efflux t
95	36	44.4	498	T04640	hypothetical prote
96	36	44.4	517	A49776	xylan 1,4-beta-xyl
97	36	44.4	524	D82220	conserved hypothet
98	36	44.4	538	H96008	hypothetical prote
99	36	44.4	558	F82236	conserved hypothet
100	36	44.4	567	D70769	hypothetical prote
101	36	44.4	575	AG2566	hypothetical prote
102	36	44.4	576	S33327	catalase (EC 1.11.

103 36 44.4 588 2 AE2573
104 36 44.4 592 2 T51712
105 36 44.4 597 2 S65673
106 36 44.4 676 1 WMBEX6
107 36 44.4 776 2 T09938
108 36 44.4 1014 2 F96501
109 36 44.4 1158 2 E86327
110 36 44.4 1257 2 S28764
111 36 44.4 1267 2 T21340
112 36 44.4 1404 1 A48196
113 36 44.4 1677 2 T18344
114 36 44.4 1680 2 T01367
115 36 44.4 1972 2 S68176
116 36 44.4 2110 2 B44110
117 36 44.4 2111 2 A70668
118 36 44.4 2118 2 T13612
119 36 44.4 2264 1 GNVVTB
120 36 44.4 3033 1 JQ1303
121 35.5 43.8 402 2 A85595
122 35.5 43.8 402 2 E90744
123 35.5 43.8 519 2 S45345
124 35.5 43.8 766 2 S45344
125 35.5 43.8 1015 2 S68141
126 35.5 43.8 1017 2 I37465
127 35 43.2 135 2 AC2305
128 35 43.2 148 2 G87639
129 35 43.2 156 1 GNVQL2
130 35 43.2 156 1 GNVQLL
131 35 43.2 156 1 GNVQMA
132 35 43.2 161 2 T22196
133 35 43.2 190 2 C89966
134 35 43.2 191 2 S41317
135 35 43.2 212 2 AC2326
136 35 43.2 220 2 I54427
137 35 43.2 233 2 C90427
138 35 43.2 246 2 I59092
139 35 43.2 258 1 DHMS
140 35 43.2 262 1 EQXL
141 35 43.2 265 2 A70640
142 35 43.2 266 2 B28031
143 35 43.2 291 1 DHMSM
144 35 43.2 302 2 S75227
145 35 43.2 308 2 AB1957
146 35 43.2 308 2 T38947
147 35 43.2 326 2 C71426
148 35 43.2 329 2 T11350
149 35 43.2 334 2 AH3282
150 35 43.2 335 2 A86464
151 35 43.2 339 2 S42595
152 35 43.2 344 2 AF3626
153 35 43.2 348 2 AB2957
154 35 43.2 348 2 D98326
155 35 43.2 357 2 T12448
156 35 43.2 364 2 A30521
157 35 43.2 379 2 I48133
158 35 43.2 379 2 I48132
159 35 43.2 379 2 I48134
160 35 43.2 382 2 T11138
161 35 43.2 384 2 S62758
162 35 43.2 407 2 T00989
163 35 43.2 414 2 S75052
164 35 43.2 417 2 D82111
165 35 43.2 420 2 A64132
166 35 43.2 421 2 A56550
167 35 43.2 427 1 T49031
168 35 43.2 443 2 H84414
169 35 43.2 447 2 T49990
170 35 43.2 447 2 F97393
171 35 43.2 477 2 B83476
172 35 43.2 501 2 E96758
173 35 43.2 532 2 A30766
174 35 43.2 536 2 T41652
175 35 43.2 536 2 AE2133

176 35 43.2 570 2 S42708
177 35 43.2 570 2 S04547
178 35 43.2 576 2 T12005
179 35 43.2 576 2 AF2361
180 35 43.2 580 2 T09808
181 35 43.2 591 1 C8HUB
182 35 43.2 635 2 A87433
183 35 43.2 659 2 T02838
184 35 43.2 676 2 D96615
185 35 43.2 694 2 S67296
186 35 43.2 699 2 D96767
187 35 43.2 716 2 S26759
188 35 43.2 762 2 H87466
189 35 43.2 819 2 T08745
190 35 43.2 819 2 T05744
191 35 43.2 838 2 G84599
192 35 43.2 879 2 S23006
193 35 43.2 886 1 RGBYL3
194 35 43.2 897 2 E69202
195 35 43.2 929 2 S75098
196 35 43.2 947 2 T01238
197 35 43.2 993 2 G96680
198 35 43.2 1015 2 B84458
199 35 43.2 1104 2 T01811
200 35 43.2 1135 2 T30561
201 35 43.2 1294 2 S58307
202 35 43.2 1307 2 T21283
203 35 43.2 1318 2 T05745
204 35 43.2 1439 2 T02087
205 35 43.2 1489 2 G71406
206 35 43.2 1505 2 S28079
207 35 43.2 1545 1 S71841
208 35 43.2 1908 2 T42707
209 35 43.2 1952 2 T48814
210 34.5 42.6 235 2 T30656
211 34.5 42.6 236 1 TVMSA1
212 34.5 42.6 236 2 I53744
213 34.5 42.6 239 1 TVHUA1
214 34.5 42.6 335 2 S58997
215 34.5 42.6 402 1 E64822
216 34.5 42.6 475 2 H84137
217 34.5 42.6 582 2 S08686
218 34.5 42.6 671 2 JE0288
219 34.5 42.6 999 1 S27756
220 34.5 42.6 1097 1 RNBYL3
221 34.5 42.6 1299 2 T43251
222 34 42.0 62 2 T13997
223 34 42.0 129 2 A82227
224 34 42.0 138 2 S34735
225 34 42.0 143 2 T16864
226 34 42.0 144 2 AC1907
227 34 42.0 147 2 JC7938
228 34 42.0 160 2 I61895
229 34 42.0 172 2 AD3606
230 34 42.0 176 1 JQ1636
231 34 42.0 204 2 F72459
232 34 42.0 220 2 A25925
233 34 42.0 240 2 S53403
234 34 42.0 244 2 E81690
235 34 42.0 247 2 AC0652
236 34 42.0 250 2 S47037
237 34 42.0 250 2 S52004
238 34 42.0 250 2 D82258
239 34 42.0 253 2 T12439
240 34 42.0 259 2 T52619
241 34 42.0 265 2 B39797
242 34 42.0 266 1 HLHU4D
243 34 42.0 266 2 A28031
244 34 42.0 269 1 EQRTA
245 34 42.0 269 2 B35678
246 34 42.0 276 2 C82420
247 34 42.0 278 2 S49623
248 34 42.0 279 2 T50748

proline transport
proline transport
NADH2 dehydrogenas
flavoprotein [impo
NADH2 dehydrogenas
complement C8 beta
hypothetical prote
probable membrane
unknown protein F1
probable membrane
unknown protein F2
pep protein - frui
beta-D-glucosidase
probable RNA helic
hypothetical prote
probable retroelem
shed acute-phase a
regulatory protein
valine-tRNA ligase
hypothetical prote
hypothetical prote
F5i14.11 imported
probable retroelem
hypothetical prote
Scythe protein - A
hypothetical 149.2
hypothetical prote
hypothetical prote
gag/pol polyprotei
probable retroviru
DNA-directed DNA p
multidrug resistan
hypothetical prote
hypothetical prote
hypothetical prote
transforming prote
gene bcl-2 protein
transforming prote
NADH2 dehydrogenas
probable membrane
hypothetical prote
finger protein zfp
krueppel-type zinc
receptor-like prot
DNA-directed DNA p
furin (EC 3.4.21.7
aquaporin - common
conserved hypothet
beta-galactosidase
hypothetical prote
hypothetical prote
type II antifreeze
eosinophil cationi
molybdopterin bios
major noncapsid pr
hypothetical prote
class II histocomp
probable membrane
probable sodium-tr
probable membrane
tonoplast intrinsi
gamma-Tip protein
cysz protein VC096
tonoplast intrinsi
TINY-like protein
MHC class II histo
MHC class II histo
MHC class II histo
enkephalin A precu
enkephalin precurs
lipase-related pro
hydroxyneurosporen

249	34	42.0	291	2	T40277	probable membrane	322	33.5	41.4	472	2	H96607	unknown protein F2
250	34	42.0	304	2	AG3035	hypothetical prote	323	33.5	41.4	548	2	B87596	xylosidase/arabino
251	34	42.0	304	2	D98250	probable sugar ABC	324	33.5	41.4	580	2	T33742	hypothetical prote
252	34	42.0	317	2	S45054	GTP-binding regula	325	33.5	41.4	619	2	S67067	probable membrane
253	34	42.0	317	2	S38398	GTP-binding protei	326	33.5	41.4	1038	1	JC5757	DNA-directed DNA p
254	34	42.0	317	2	A33928	GTP-binding protei	327	33.5	41.4	1038	2	T18222	DNA polymerase del
255	34	42.0	317	2	B33928	GTP-binding protei	328	33.5	41.4	1247	2	T45743	hypothetical prote
256	34	42.0	317	2	A36986	activated protein	329	33.5	41.4	1496	2	AH0447	insecticidal toxin
257	34	42.0	324	2	T01316	epoxide hydrolase	330	33.5	41.4	2061	2	T13751	transcription fact
258	34	42.0	329	2	T07000	chitinase (EC 3.2.	331	33.5	41.4	2318	2	S45306	notch 3 protein -
259	34	42.0	336	2	A46478	major surface anti	332	33.5	41.4	4574	2	G02520	plectin [imported]
260	34	42.0	336	2	A30527	major surface anti	333	33.5	41.4	4684	2	A59404	ig heavy chain V r
261	34	42.0	343	2	S62704	ubiquinol-cytochro	334	33	40.7	32	2	S20771	ATP synthase F0 ch
262	34	42.0	357	2	T11703	NADH2 dehydrogenas	335	33	40.7	55	2	E90618	hypothetical prote
263	34	42.0	360	2	AE0068	phospho-N-acetylm	336	33	40.7	88	2	H91290	hypothetical prote
264	34	42.0	370	1	B43673	chloromuconate cyc	337	33	40.7	103	2	S59846	hypothetical prote
265	34	42.0	377	2	F95982	probable transport	338	33	40.7	106	2	D69808	hypothetical prote
266	34	42.0	393	2	D95894	probable hydrolase	339	33	40.7	112	2	F75079	hypothetical prote
267	34	42.0	394	2	S75996	hypothetical prote	340	33	40.7	128	2	T22276	hypothetical prote
268	34	42.0	396	2	T23619	hypothetical prote	341	33	40.7	138	2	E65019	hypothetical prote
269	34	42.0	430	2	B70959	probable dihydrooor	342	33	40.7	142	2	AC1934	hypothetical prote
270	34	42.0	435	1	WZBE50	gene 50 protein -	343	33	40.7	149	2	S74366	hypothetical prote
271	34	42.0	449	2	F81954	ubiquinol-cytochro	344	33	40.7	156	2	AG2741	conserved hypothet
272	34	42.0	449	2	E81011	ubiquinol-cytochro	345	33	40.7	164	2	AE1834	hypothetical prote
273	34	42.0	449	2	C81817	probable integral	346	33	40.7	166	2	I39546	hypothetical prote
274	34	42.0	477	2	H83588	probable MFS trans	347	33	40.7	172	2	F97522	hypothetical prote
275	34	42.0	478	2	E86314	F2H15.15 protein -	348	33	40.7	185	2	I40478	conserved hypothet
276	34	42.0	483	2	T10552	hypothetical prote	349	33	40.7	194	2	H85097	hypothetical prote
277	34	42.0	486	2	F86654	transporter ycdH [350	33	40.7	202	2	G85439	hypothetical prote
278	34	42.0	503	2	T05347	hypothetical prote	351	33	40.7	231	2	S60769	hypothetical prote
279	34	42.0	511	1	E70391	major facilitator	352	33	40.7	231	2	AF3380	DNA repair/recombi
280	34	42.0	532	1	E69821	multidrug resistan	353	33	40.7	232	2	C70699	hypothetical prote
281	34	42.0	535	2	B84239	hypothetical prote	354	33	40.7	238	2	T28337	probable paba prot
282	34	42.0	561	2	AC0605	probable membrane	355	33	40.7	240	1	JJAGTT	probable metallopr
283	34	42.0	561	2	G90744	probable transport	356	33	40.7	240	2	T12191	adenylate isopente
284	34	42.0	561	2	C85595	probable membrane	357	33	40.7	240	2	AB1871	NADH2 dehydrogenas
285	34	42.0	561	2	G64822	TrkA, potassium ch	358	33	40.7	246	2	D83540	hypothetical prote
286	34	42.0	562	2	AI0161	hypothetical prote	359	33	40.7	247	2	B84962	probable sulfate u
287	34	42.0	566	2	T15226	6fe-6S prismane cl	360	33	40.7	249	2	A70840	hypothetical prote
288	34	42.0	567	2	C97321	hypothetical prote	361	33	40.7	252	2	T01947	hypothetical prote
289	34	42.0	574	2	AH2292	probable exported	362	33	40.7	253	2	TI4518	probable membrane
290	34	42.0	618	2	AI0171	traG protein - Esc	363	33	40.7	256	2	JQ1106	hypothetical prote
291	34	42.0	635	2	S22999	traG protein - Ent	364	33	40.7	256	2	S26742	tonoplast intrinsi
292	34	42.0	637	2	T08530	traG protein - Esc	365	33	40.7	264	2	AH0759	tonoplast intrinsi
293	34	42.0	637	2	S22992	probable drug tran	366	33	40.7	269	2	T10253	propanediol diffus
294	34	42.0	643	2	G75599	hypothetical prote	367	33	40.7	275	2	H70706	membrane protein M
295	34	42.0	687	2	AG2385	beta-D-glucosidase	368	33	40.7	279	2	T10251	probable OXIDOREDU
296	34	42.0	758	2	D87369	hypothetical prote	369	33	40.7	281	2	S04405	membrane protein M
297	34	42.0	766	2	T28120	hypothetical prote	370	33	40.7	282	2	S61663	hydroxyneurosporen
298	34	42.0	803	2	T39530	hypothetical prote	371	33	40.7	284	2	B82955	probable membrane
299	34	42.0	884	2	S77031	hypothetical prote	372	33	40.7	289	2	AB0262	hypothetical prote
300	34	42.0	940	2	T01854	hypothetical prote	373	33	40.7	290	2	T23268	conserved hypothet
301	34	42.0	974	2	A72012	metalloproteinase,	374	33	40.7	306	2	T26639	hypothetical prote
302	34	42.0	974	2	B86613	zinc metalloprotei	375	33	40.7	308	1	QQLJX1	trans-activating t
303	34	42.0	975	2	T29908	hypothetical prote	376	33	40.7	311	2	TI1362	NADH2 dehydrogenas
304	34	42.0	985	2	I51672	receptor tyrosine	377	33	40.7	320	2	AI2305	hypothetical prote
305	34	42.0	1008	2	D84434	probable receptor-	378	33	40.7	322	2	AG3201	transcription regu
306	34	42.0	1025	2	A82516	type I restriction	379	33	40.7	324	2	G82410	transcription regu
307	34	42.0	1064	2	A41542	adenylate cyclase	380	33	40.7	325	2	D70728	hypothetical prote
308	34	42.0	1179	2	T05673	hypothetical prote	381	33	40.7	326	2	AH2612	NADH-ubiquinone ox
309	34	42.0	1404	2	F86470	probable retroelem	382	33	40.7	328	2	I47161	ig gamma 3 chain c
310	34	42.0	1413	2	G86301	probable retroelem	383	33	40.7	328	2	I47158	ig gamma 1 chain c
311	34	42.0	1435	1	BVBYL1	guanine nucleotide	384	33	40.7	329	2	T40126	hypothetical prote
312	34	42.0	1468	2	T05672	hypothetical prote	385	33	40.7	340	2	S49742	hypothetical prote
313	34	42.0	1567	2	T03730	antigen containing	386	33	40.7	342	2	E64905	probable sugar tra
314	34	42.0	1777	2	AC2088	serine/threonine k	387	33	40.7	342	2	A90894	probable transport
315	34	42.0	2476	2	T34022	zonadhesin - pig	388	33	40.7	342	2	H85723	probable transport
316	34	42.0	2476	2	TL3931	projectin - fruit	389	33	40.7	343	2	A35639	G protein-coupled
317	34	41.4	236	2	JC7383	B-cell lymphoma 2	390	33	40.7	345	2	G86372	protein F508.39 [i
318	33.5	41.4	245	2	C86418	hypothetical prote	391	33	40.7	346	2	I48185	gene alx3 protein
319	33.5	41.4	343	2	B97256	D-alanine-D-alanin	392	33	40.7	349	2	AD3650	trans-1,2-dihydrob
320	33.5	41.4	353	2	AF2053	hypothetical prote	393	33	40.7	352	2	B69901	fatty-acid desatur
321	33.5	41.4	399	2	AC1601	hypothetical prote	394	33	40.7	354	2	F97394	probable NADH-ubiq

395	33	40.7	355	2	C64644	integrase/recombin	468	33	40.7	578	2	T48795	origin recognition
396	33	40.7	361	2	A97837	queuine tRNA-ribos	469	33	40.7	582	2	H70588	probable pros prot
397	33	40.7	362	2	G69589	3-dehydroquinat s	470	33	40.7	594	2	S37816	hypothetical prote
398	33	40.7	367	2	T06706	hypothetical prote	471	33	40.7	612	2	T09867	NADH2 dehydrogenas
399	33	40.7	370	2	T44667	chloromuconate cyc	472	33	40.7	614	2	A41757	betaine transport
400	33	40.7	370	2	T44618	chloromuconate cyc	473	33	40.7	622	2	AG0952	membrane transport
401	33	40.7	370	2	JQ0176	chloromuconate cyc	474	33	40.7	622	2	A98215	low affinity potas
402	33	40.7	371	2	B86489	protein T32E20.25	475	33	40.7	622	2	A49345	K(+)-uptake protei
403	33	40.7	379	2	I48135	ubiquinol-cytochro	476	33	40.7	622	2	B86061	low affinity potas
404	33	40.7	379	2	I48180	ubiquinol-cytochro	477	33	40.7	624	2	T49366	myocyte-specific e
405	33	40.7	380	2	A71390	ubiquinol-cytochro	478	33	40.7	635	2	AE2086	hypothetical prote
406	33	40.7	382	2	D58930	ubiquinol-cytochro	479	33	40.7	653	2	A49722	endoglin precursor
407	33	40.7	399	2	T51472	hypothetical prote	480	33	40.7	655	1	A46688	hepatocyte growth
408	33	40.7	400	2	A41679	beta-3-adrenergic	481	33	40.7	658	2	S50831	endoglin precursor
409	33	40.7	400	2	A53281	beta 3-adrenergic	482	33	40.7	660	2	G82672	ATP sulfurylase, l
410	33	40.7	400	2	T49810	hypothetical prote	483	33	40.7	662	2	T44219	hypothetical prote
411	33	40.7	402	2	A70882	probable pps prote	484	33	40.7	662	2	T44034	hypothetical prote
412	33	40.7	405	2	T44249	transport protein	485	33	40.7	687	2	B70515	hypothetical prote
413	33	40.7	407	2	T36404	probable monooxyge	486	33	40.7	688	2	T05353	hypothetical prote
414	33	40.7	409	2	T08928	sucrose cleavage p	487	33	40.7	714	2	T14080	hypothetical prote
415	33	40.7	412	2	G87522	hypothetical prote	488	33	40.7	756	2	T34797	probable ATP bindi
416	33	40.7	415	2	B97512	probable cytochrom	489	33	40.7	764	2	T01493	probable potassium
417	33	40.7	415	2	AC2731	cytochrome P450 hy	490	33	40.7	780	2	H84685	probable vacuolar
418	33	40.7	416	2	A37877	hemorrhagic protei	491	33	40.7	795	2	T34673	probable SecDF pro
419	33	40.7	418	2	E83686	hypothetical prote	492	33	40.7	856	2	T22575	hypothetical prote
420	33	40.7	422	2	G83503	probable MFS trans	493	33	40.7	873	2	T12535	hypothetical prote
421	33	40.7	422	2	A71058	probable cytosine	494	33	40.7	894	2	B96557	probable receptor
422	33	40.7	427	2	S30558	ribose-phosphate d	495	33	40.7	899	2	T07062	probable lipoxigen
423	33	40.7	427	2	C86469	protein F12K21.14	496	33	40.7	911	2	AD2271	cation-transportin
424	33	40.7	435	2	T29593	hypothetical prote	497	33	40.7	992	2	T38817	hypothetical prote
425	33	40.7	436	1	S49458	diphosphate-fructo	498	33	40.7	1007	2	T24643	hypothetical prote
426	33	40.7	444	2	F87366	cytochrome P450 fa	499	33	40.7	1070	1	A54600	1-phosphatidylinos
427	33	40.7	448	2	C83347	probable dipeptida	500	33	40.7	1076	2	T30842	serine-repeat anti
428	33	40.7	449	2	F83328	probable sodium/al	501	33	40.7	1088	1	PFR7GA	platelet-derived g
429	33	40.7	451	2	A55909	transforming prote	502	33	40.7	1100	2	T42260	guanylate cyclase
430	33	40.7	456	1	I40516	spaF protein - Bac	503	33	40.7	1121	2	JC7329	WD-repeat protein
431	33	40.7	456	2	E83750	gluconate permease	504	33	40.7	1194	1	G70837	probable ABC trans
432	33	40.7	456	2	A83264	hypothetical prote	505	33	40.7	1311	2	G86471	unknown protein [i
433	33	40.7	462	2	AD1284	menaquinone-specif	506	33	40.7	1385	2	T13415	hypothetical prote
434	33	40.7	462	2	AG1655	menaquinone-specif	507	33	40.7	1415	2	T08945	hypothetical prote
435	33	40.7	464	2	T16889	hypothetical prote	508	33	40.7	1433	2	A71444	probable LTR retro
436	33	40.7	466	2	H87419	conserved hypothet	509	33	40.7	1454	2	E84535	probable retroelem
437	33	40.7	469	2	I37451	HBFG-2 (HFK-2) pro	510	33	40.7	1461	2	E84589	probable retroelem
438	33	40.7	476	2	A54743	transcription fact	511	33	40.7	1474	2	D88550	protein ZC84.6 [im
439	33	40.7	479	2	S73770	NADH oxidase nox -	512	33	40.7	1502	2	D84587	probable myosin he
440	33	40.7	480	2	JH0672	brain factor 1 pro	513	33	40.7	1522	2	C96578	hypothetical prote
441	33	40.7	482	2	A96790	protein T23E18.18	514	33	40.7	1541	1	S71839	canalicular multid
442	33	40.7	489	2	A11276	multidrug-efflux t	515	33	40.7	1697	2	T00079	hypothetical prote
443	33	40.7	489	2	A11639	multidrug-efflux t	516	33	40.7	2049	2	T29227	hypothetical prote
444	33	40.7	489	2	G71245	hypothetical prote	517	33	40.7	2210	1	RRXPTV	genome polypeptide
445	33	40.7	491	2	T48989	hypothetical prote	518	33	40.7	2824	2	T22759	hypothetical prote
446	33	40.7	493	2	AD1398	drug-export protei	519	33	40.7	2844	2	S28291	hypothetical prote
447	33	40.7	493	2	AG1773	drug-export protei	520	33	40.7	3033	1	GNWVJ8	genome polypeptide
448	33	40.7	496	2	T00978	probable pectinest	521	33	40.7	4344	1	A53489	dynein heavy chain
449	33	40.7	498	2	JC8008	beta3-glycosyltran	522	33	40.7	4464	2	D87755	protein T21E12.4 [
450	33	40.7	502	2	C86263	hypothetical prote	523	33	40.7	4687	1	A39638	plectin - rat
451	33	40.7	506	2	B87102	conserved membrane	524	33	40.7	4725	1	A44357	dynein heavy chain
452	33	40.7	513	2	T14864	probable monosacch	525	33	40.7	5376	2	T42215	zonadhesin - mouse
453	33	40.7	516	2	D82279	sodium/alanine sym	526	32.5	40.1	75	2	AC2441	hypothetical prote
454	33	40.7	519	2	D65178	kup protein - Bach	527	32.5	40.1	122	4	JN0029	serum amyloid A3 p
455	33	40.7	525	2	T47409	hypothetical prote	528	32.5	40.1	140	2	D96732	hypothetical prote
456	33	40.7	537	2	B46535	interleukin 2 rece	529	32.5	40.1	174	2	T38498	hypothetical prote
457	33	40.7	549	2	S48169	metalloproteinase	530	32.5	40.1	359	1	B49107	hypothetical prote
458	33	40.7	549	2	E86337	hypothetical prote	531	32.5	40.1	375	1	B49107	alcohol dehydrogen
459	33	40.7	552	2	T11161	NADH2 dehydrogenas	532	32.5	40.1	401	2	AG1238	hypothetical prote
460	33	40.7	567	2	C85643	hypothetical prote	533	32.5	40.1	403	2	B46165	envelope surface g
461	33	40.7	567	2	H90782	hypothetical prote	534	32.5	40.1	442	2	D84224	heme synthase limpo
462	33	40.7	572	2	T11478	NADH2 dehydrogenas	535	32.5	40.1	528	2	T43919	yfuB protein limpo
463	33	40.7	572	2	T16865	hypothetical prote	536	32.5	40.1	654	2	AB2932	beta 1,3 glucan sy
464	33	40.7	573	2	H25797	NADH2 dehydrogenas	537	32.5	40.1	654	2	D98350	hypothetical prote
465	33	40.7	574	2	S01186	NADH2 dehydrogenas	538	32.5	40.1	869	2	A55384	transcription fact
466	33	40.7	576	1	HNNZS	hemagglutinin-neur	539	32.5	40.1	873	2	B75514	penicillin-binding
467	33	40.7	576	2	H82777	alkaline phosphata	540	32.5	40.1	968	2	G81743	preprotein translo

541	32.5	40.1	969	2	G71482	probable protein t	614	32	39.5	308	2	F95903	conserved hypothet
542	32.5	40.1	970	2	A72028	preprotein translo	615	32	39.5	311	2	T32118	hypothetical prote
543	32.5	40.1	970	2	G86595	protein translocas	616	32	39.5	321	2	T11274	NADH2 dehydrogenas
544	32.5	40.1	1986	2	S28353	probable polyketid	617	32	39.5	331	2	T35100	probable secreted
545	32	39.5	53	2	T31169	hypothetical prote	618	32	39.5	335	2	T48319	hypothetical prote
546	32	39.5	61	2	T47055	hypothetical prote	619	32	39.5	338	2	A44132	D-aspartate oxidas
547	32	39.5	66	2	S09855	hypothetical prote	620	32	39.5	339	2	H83252	glycosyltransferas
548	32	39.5	87	2	S15699	hypothetical prote	621	32	39.5	340	2	H91084	probable glucarate
549	32	39.5	93	2	T17447	hypothetical prote	622	32	39.5	341	2	S57595	mitochondrial repl
550	32	39.5	93	2	AC0232	probable transcrip	623	32	39.5	342	2	JC7110	brain-specific mem
551	32	39.5	102	2	S69752	hypothetical prote	624	32	39.5	343	2	A41748	lumican precursor
552	32	39.5	102	2	T49651	hypothetical prote	625	32	39.5	351	1	VCVETC	coat protein - tur
553	32	39.5	113	2	S26277	hypothetical prote	626	32	39.5	354	2	G71465	hypothetical prote
554	32	39.5	113	2	I38318	T-cell receptor be	627	32	39.5	354	2	A81728	conserved hypothet
555	32	39.5	113	2	S26276	T-cell receptor be	628	32	39.5	360	2	S00834	int-1-like protein
556	32	39.5	115	2	C83574	T-cell receptor be	629	32	39.5	360	2	B36470	Wnt-2 protein - mo
557	32	39.5	120	2	S06731	hypothetical prote	630	32	39.5	361	2	T40857	hypothetical prote
558	32	39.5	121	2	S29002	Ig kappa chain pre	631	32	39.5	362	2	E96587	hypothetical prote
559	32	39.5	125	2	F81197	trypsin inhibitor	632	32	39.5	364	2	AF0421	probable sugar ABC
560	32	39.5	132	2	C97738	holo-(acyl-carrier	633	32	39.5	365	2	A85930	probable glucarate
561	32	39.5	134	2	B75468	hypothetical prote	634	32	39.5	367	2	S59146	ubiquinol-cytochro
562	32	39.5	136	2	G71601	hypothetical prote	635	32	39.5	368	2	T04532	probable (S)-2-hyd
563	32	39.5	143	2	H72610	hypothetical prote	636	32	39.5	368	2	G85206	glycolate oxidase-
564	32	39.5	144	2	S38391	hypothetical prote	637	32	39.5	371	2	C87446	queuine tRNA ribos
565	32	39.5	147	2	D69164	T-cell receptor be	638	32	39.5	371	2	G87644	hypothetical prote
566	32	39.5	150	1	W6WL6	hypothetical prote	639	32	39.5	379	2	AG0585	citrate utilizatio
567	32	39.5	151	2	A97008	E6 protein - human	640	32	39.5	383	1	DHHU	Ig delta chain C r
568	32	39.5	157	2	T34552	transcription regu	641	32	39.5	384	2	T11931	ubiquinol-cytochro
569	32	39.5	161	2	G81544	hypothetical prote	642	32	39.5	385	2	F70604	probable fadE30 pr
570	32	39.5	162	2	C98011	conserved hypothet	643	32	39.5	387	2	AI0511	probable metabolit
571	32	39.5	172	2	T34785	conserved hypothet	644	32	39.5	387	2	E97028	probable amidohydr
572	32	39.5	172	2	AH0462	hypothetical prote	645	32	39.5	392	2	AI3141	Xaa-Pro dipeptidas
573	32	39.5	182	2	D65005	probable exported	646	32	39.5	392	2	C98146	Xaa-Pro dipeptidas
574	32	39.5	182	2	C85874	hypothetical prote	647	32	39.5	395	2	B96610	hypothetical prote
575	32	39.5	182	2	B91030	probable transport	648	32	39.5	397	2	T35713	probable oxidoredu
576	32	39.5	182	2	G70687	probable transport	649	32	39.5	399	2	F83796	multidrug-efflux t
577	32	39.5	196	2	S45553	hypothetical prote	650	32	39.5	400	2	T05563	hypothetical prote
578	32	39.5	197	2	JC2320	spore maturation p	651	32	39.5	401	2	AD1264	ammonium transport
579	32	39.5	209	2	S27494	hypothetical 22K p	652	32	39.5	403	1	GRECY	tyrosine-specific
580	32	39.5	214	2	B40296	nods protein - Bra	653	32	39.5	403	2	G90955	tyrosine-specific
581	32	39.5	222	2	S63320	hypothetical prote	654	32	39.5	403	2	D85804	tyrosine-specific
582	32	39.5	222	2	S56195	probable membrane	655	32	39.5	403	2	A10604	probable membrane
583	32	39.5	222	2	AC2397	ATP-binding protei	656	32	39.5	405	2	A75459	sodium extrusion p
584	32	39.5	225	2	T26491	hypothetical prote	657	32	39.5	412	2	F91068	hypothetical prote
585	32	39.5	226	2	T08335	hypothetical prote	658	32	39.5	412	2	H85912	hypothetical prote
586	32	39.5	233	2	T47136	hypothetical prote	659	32	39.5	413	2	B90162	DNA/pantothenate m
587	32	39.5	238	2	H83371	probable amino aci	660	32	39.5	413	2	T28646	Y4jS protein - Rhi
588	32	39.5	238	2	AD2034	hypothetical prote	661	32	39.5	414	2	AI2295	two-component resp
589	32	39.5	245	2	C95314	TRm23b IS ATP-bind	662	32	39.5	415	2	T46716	hypothetical prote
590	32	39.5	245	2	D95334	TRm23b IS ATP-bind	663	32	39.5	416	2	C81324	ubiquinol-cytochro
591	32	39.5	256	1	S10164	fumarate reductase	664	32	39.5	419	2	A83133	probable MFS trans
592	32	39.5	261	2	C87591	hydrolase, alpha/b	665	32	39.5	419	2	T24820	hypothetical prote
593	32	39.5	262	1	S17705	indole-3-glycerol-	666	32	39.5	420	2	D83952	hypothetical prote
594	32	39.5	263	1	S43189	hypothetical prote	667	32	39.5	421	2	T27311	hypothetical prote
595	32	39.5	265	2	H84867	probable endochiti	668	32	39.5	426	2	H87307	ubiquinol-cytochro
596	32	39.5	265	2	G72126	ct338 hypothetical	669	32	39.5	432	2	A70714	hypothetical prote
597	32	39.5	265	2	F86495	CT338 hypothetical	670	32	39.5	438	2	AD1888	hypothetical prote
598	32	39.5	265	2	A86548	hypothetical prote	671	32	39.5	438	2	T51395	hypothetical prote
599	32	39.5	265	2	H72075	hypothetical prote	672	32	39.5	439	2	T01270	hypothetical prote
600	32	39.5	265	2	C81592	hypothetical prote	673	32	39.5	445	2	D82342	probable cadaverin
601	32	39.5	266	2	T16106	hypothetical prote	674	32	39.5	446	2	AD0861	glucarate dehydrat
602	32	39.5	268	2	I53029	enkephalin A precu	675	32	39.5	446	2	H65060	glucarate dehydrat
603	32	39.5	268	2	B82929	hypothetical prote	676	32	39.5	447	2	AE0094	probable gluconate
604	32	39.5	272	2	C96742	unknown protein F1	677	32	39.5	453	2	T48240	hypothetical prote
605	32	39.5	274	2	C83709	hypothetical prote	678	32	39.5	456	2	AE0164	probable exported
606	32	39.5	285	2	B75507	hypothetical prote	679	32	39.5	458	2	T47389	hypothetical prote
607	32	39.5	286	2	T41931	minor capsid prote	680	32	39.5	459	2	T45559	NADH2 dehydrogenas
608	32	39.5	290	2	T16540	hypothetical prote	681	32	39.5	460	2	G01210	guanine nucleotide
609	32	39.5	293	2	AB2295	succinyl-CoA synth	682	32	39.5	463	2	A44808	cellulase (EC 3.2.
610	32	39.5	295	2	AE1019	lysR family regula	683	32	39.5	465	2	G90961	flagellar capping
611	32	39.5	297	2	AB1962	acetylglutamate ki	684	32	39.5	465	2	G85809	flagellar capping
612	32	39.5	303	2	S73051	tropinesterase hom	685	32	39.5	468	2	A64956	flagellar hook-ass
613	32	39.5	304	2	H86332	T20H2.26 protein -	686	32	39.5	469	2	AE0835	probable type I se

687 32 39.5 469 2 A87467 glycosyl hydrolase
688 32 39.5 471 2 AG0146 probable pyridoxal
689 32 39.5 471 2 T32788 hypothetical prote
690 32 39.5 478 2 F82175 conserved hypothet
691 32 39.5 484 2 AE0887 probable membrane
692 32 39.5 490 2 F87443 conserved hypothet
693 32 39.5 492 2 T28025 hypothetical prote
694 32 39.5 494 2 AG3103 mannitol 2-dehydro
695 32 39.5 494 2 A82294 probable carbon st
696 32 39.5 496 2 A54757 protein disulfide-
697 32 39.5 500 2 T36090 probable integral
698 32 39.5 500 2 A86112 probable transport
699 32 39.5 503 2 E83490 probable MFS trans
700 32 39.5 505 2 E87021 probable integral-
701 32 39.5 506 2 H70928 hypothetical prote
702 32 39.5 509 2 S17597 Ig delta chain (WI
703 32 39.5 509 2 A82840 beta-lactamase ind
704 32 39.5 509 2 E83002 drug efflux transp
705 32 39.5 510 2 H82624 di-tripeptide ABC
706 32 39.5 511 2 C69199 phenylalanine-tRNA
707 32 39.5 514 2 S56384 hypothetical 56.3K
708 32 39.5 514 2 A91271 probable transport
709 32 39.5 516 2 H70935 hypothetical prote
710 32 39.5 518 2 AD1930 ammonium transport
711 32 39.5 520 2 T21462 hypothetical prote
712 32 39.5 522 2 S77073 hypothetical prote
713 32 39.5 525 2 C98183 mannitol 2-dehydro
714 32 39.5 539 2 A35052 interleukin-2 rece
715 32 39.5 544 2 F81059 YhbX/YhjW/YijP/Yjd
716 32 39.5 547 2 T39478 zinc-finger protei
717 32 39.5 557 2 F69481 probable acid-CoA
718 32 39.5 564 2 AG2823 ABC transporter, m
719 32 39.5 564 2 F97601 afub (AE006182) [i
720 32 39.5 583 2 T20291 hypothetical prote
721 32 39.5 598 2 E83583 probable acyl-CoA
722 32 39.5 598 2 T51033 hypothetical prote
723 32 39.5 602 2 A44409 gamma-aminobutyric
724 32 39.5 602 2 A45078 gamma-aminobutyric
725 32 39.5 615 2 AI0615 probable exported
726 32 39.5 624 2 S50650 hypothetical prote
727 32 39.5 640 2 T28631 Y4CD protein - Rhi
728 32 39.5 650 2 JC7095 sodium-dependent v
729 32 39.5 653 2 A84675 hypothetical prote
730 32 39.5 657 2 C84675 hypothetical prote
731 32 39.5 662 2 T17211 hypothetical prote
732 32 39.5 667 2 T09013 RING finger protei
733 32 39.5 698 2 H75577 probable drug tran
734 32 39.5 703 2 G69186 ribonucleoside-dip
735 32 39.5 705 2 D84680 probable ABC trans
736 32 39.5 707 2 T24361 hypothetical prote
737 32 39.5 708 1 A32959 tryptophan synthas
738 32 39.5 720 2 E86297 F3O9.6 protein - A
739 32 39.5 742 2 S27263 synaptic vesicle p
740 32 39.5 755 1 S74695 hypothetical prote
741 32 39.5 755 2 T46411 hypothetical prote
742 32 39.5 757 2 B84790 hypothetical prote
743 32 39.5 771 2 S72526 inorganic diphosph
744 32 39.5 776 1 RGECAE aerobic respiratio
745 32 39.5 778 2 AG0906 aerobic respiratio
746 32 39.5 778 2 A91140 aerobic respiratio
747 32 39.5 778 2 D85985 aerobic respiratio
748 32 39.5 781 2 A85035 hypothetical prote
749 32 39.5 781 2 T51433 probable cation tr
750 32 39.5 784 2 AH2560 hypothetical prote
751 32 39.5 786 2 D75630 glycerophosphoryl
752 32 39.5 788 2 AF0122 probable membrane
753 32 39.5 792 2 G84830 probable potassium
754 32 39.5 796 2 A45695 capsid precursor -
755 32 39.5 797 2 E84642 hypothetical prote
756 32 39.5 806 2 S76650 sensory transducti
757 32 39.5 846 2 AI2336 primosomal protein
758 32 39.5 869 2 AD2418 SWI/SNF family hel
759 32 39.5 896 2 T07408 lipoxigenase (EC 1

104K glycoprotein
hypothetical prote
probable zinc meta
metalloprotease,
receptor protein-t
isoleucyl-tRNA syn
hypothetical prote
platelet-derived g
probable arabinosy
probable Arpase F2
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
unknown protein [i
DNA polymerase hom
DNA polymerase - c
disease resistance
hypothetical prote
neurocan - mouse
hypothetical prote
hypothetical prote
hypothetical prote
collagen alpha 2(I
guanine nucleotide
FYVE finger-contai
ribulose biphosph
genome polyprotein
hypothetical prote
genome polyprotein
surface protein ty
notch protein - fr
hypothetical prote
hypothetical prote
BPLF1 protein - hu
delta-(L-alpha-ami
hypothetical coile
T-cell receptor al
Ig heavy chain V r
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
conserved hypothet
T-cell receptor be
hypothetical prote
hypothetical prote
UDP-3-0-acyl N-ace
dUTP diphosphatase
glycine betaine/l-
hypothetical prote
dihydrodipicolinat
probable aspartoki
aspartokinase III
rhamnogalacturonas
nitric-oxide reduc
probable betaine-a
apolipoprotein N-a
probable aldehyde-
probable sucrose t
env polyprotein pr
hypothetical prote
probable lipopolys
lipopolysaccharide
glutamine transpor
probable DEAH ATP-
reverse transcript
melittin, minor -
hypothetical prote

687 760 32 39.5 901 1 WMNVTN
688 761 32 39.5 914 2 S18942
689 762 32 39.5 974 2 A71466
690 763 32 39.5 975 2 C81728
691 764 32 39.5 986 2 I78844
692 765 32 39.5 1059 2 E87058
693 766 32 39.5 1071 2 A84785
694 767 32 39.5 1089 1 S33727
695 768 32 39.5 1098 2 G70697
696 769 32 39.5 1123 2 C96622
697 770 32 39.5 1139 2 T23018
698 771 32 39.5 1141 2 T05068
699 772 32 39.5 1200 2 F96711
700 773 32 39.5 1203 2 T04294
701 774 32 39.5 1213 2 D86384
702 775 32 39.5 1214 2 T18549
703 776 32 39.5 1214 2 T30941
704 777 32 39.5 1239 2 T06143
705 778 32 39.5 1244 2 C96584
706 779 32 39.5 1268 2 S52781
707 780 32 39.5 1315 2 G96722
708 781 32 39.5 1340 2 AH2219
709 782 32 39.5 1633 2 T01879
710 783 32 39.5 1712 1 CGHU2B
711 784 32 39.5 1793 2 T47897
712 785 32 39.5 2052 2 T18290
713 786 32 39.5 2149 2 C96695
714 787 32 39.5 2210 1 RRXPPLC
715 788 32 39.5 2241 2 S09811
716 789 32 39.5 2252 2 S06188
717 790 32 39.5 2395 1 S50820
718 791 32 39.5 2703 1 A24420
719 792 32 39.5 2948 2 T22664
720 793 32 39.5 3083 2 AH2493
721 794 32 39.5 3149 1 QOBE8
722 795 32 39.5 3649 1 S18268
723 796 32 39.5 4717 2 T41581
724 797 31.5 38.9 32 2 S21547
725 798 31.5 38.9 105 2 PH1020
726 799 31.5 38.9 113 2 S26263
727 800 31.5 38.9 113 2 S26262
728 801 31.5 38.9 113 2 S17385
729 802 31.5 38.9 113 2 S26266
730 803 31.5 38.9 113 2 S17386
731 804 31.5 38.9 122 2 A83756
732 805 31.5 38.9 146 2 S26408
733 806 31.5 38.9 148 2 D87464
734 807 31.5 38.9 152 2 S21826
735 808 31.5 38.9 175 2 H38164
736 809 31.5 38.9 182 2 T49852
737 810 31.5 38.9 274 2 T04619
738 811 31.5 38.9 276 2 S74967
739 812 31.5 38.9 278 1 QOBE17
740 813 31.5 38.9 278 2 AB3307
741 814 31.5 38.9 352 2 T22384
742 815 31.5 38.9 380 1 WZZMP
743 816 31.5 38.9 431 2 G71524
744 817 31.5 38.9 437 2 D81681
745 818 31.5 38.9 440 2 A55415
746 819 31.5 38.9 466 2 S53713
747 820 31.5 38.9 505 2 T03394
748 821 31.5 38.9 505 2 A82261
749 822 31.5 38.9 522 2 T50359
750 823 31.5 38.9 537 2 T02982
751 824 31.5 38.9 584 1 VCM5IA
752 825 31.5 38.9 612 2 T13616
753 826 31.5 38.9 622 2 G81981
754 827 31.5 38.9 623 2 B81037
755 828 31.5 38.9 663 2 S69566
756 829 31.5 38.9 965 2 T17395
757 830 31.5 38.9 1258 2 T14855
758 831 31 38.3 27 1 MEHB2
759 832 31 38.3 39 2 E82725

833 31 38.3 46 2 T37121 hypothetical prote
834 31 38.3 59 2 C91211 hypothetical prote
835 31 38.3 63 2 T08130 oleosin-like prote
836 31 38.3 86 2 G95993 hypothetical prote
837 31 38.3 89 2 C42525 A-ORF-P prote in -
838 31 38.3 89 2 JQ1822 SalFa prote in - va
839 31 38.3 98 2 G84976 factor-for-inversi
840 31 38.3 99 2 D75488 conserved hypothet
841 31 38.3 106 4 S57386 hypothetical prote
842 31 38.3 107 2 PL0270 Ig kappa chain V r
843 31 38.3 107 2 PL0272 Ig kappa chain V r
844 31 38.3 107 2 PL0271 Ig kappa chain V r
845 31 38.3 107 2 PL0269 Ig kappa chain V r
846 31 38.3 109 2 S74284 hypothetical prote
847 31 38.3 110 2 S13688 Ig heavy chain V r
848 31 38.3 111 2 S13687 Ig heavy chain V r
849 31 38.3 112 2 S13686 Ig heavy chain V r
850 31 38.3 112 2 S13685 Ig heavy chain V r
851 31 38.3 112 2 S51123 genome polyprotein
852 31 38.3 122 4 T01781 probable pol prote
853 31 38.3 123 2 G82357 conserved hypothet
854 31 38.3 124 2 B27664 T-cell receptor be
855 31 38.3 124 2 T32823 hypothetical prote
856 31 38.3 127 2 C75315 diacylglycerol kin
857 31 38.3 130 2 AF1885 hypothetical prote
858 31 38.3 130 2 A26316 interferon-induced
859 31 38.3 133 2 F81783 hypothetical prote
860 31 38.3 133 2 D81174 hypothetical prote
861 31 38.3 134 1 WTBO seminal fluid prot
862 31 38.3 140 2 S54245 Ig mu heavy chain
863 31 38.3 140 2 G83187 conserved hypothet
864 31 38.3 141 2 G27577 T-cell receptor be
865 31 38.3 142 2 A27608 T-cell receptor be
866 31 38.3 143 2 E96638 hypothetical prote
867 31 38.3 146 2 E70557 hypothetical prote
868 31 38.3 147 2 T04470 probable trypsin i
869 31 38.3 153 2 D24372 dpnD prote in - Str
870 31 38.3 153 2 F98079 restriction system
871 31 38.3 153 2 H95215 DpnD prote in limpo
872 31 38.3 160 2 H72569 hypothetical prote
873 31 38.3 163 1 E69231 conserved hypothet
874 31 38.3 167 2 T45261 conserved hypothet
875 31 38.3 169 2 B26395 T-cell receptor be
876 31 38.3 173 2 H87629 conserved hypothet
877 31 38.3 173 2 G96005 probable RNA polym
878 31 38.3 174 2 A49181 alpha B-crystallin
879 31 38.3 174 2 S74666 DnaJ prote in - Syn
880 31 38.3 176 2 S45343 glomerulosclerosis
881 31 38.3 176 2 S29031 mpv17 prote in - mo
882 31 38.3 177 2 H87412 hypothetical prote
883 31 38.3 178 2 JQ1547 stripe disease-spe
884 31 38.3 181 2 A97878 phosphoribosylglyc
885 31 38.3 181 2 D95005 phosphoribosylglyc
886 31 38.3 182 2 C84982 hypothetical prote
887 31 38.3 188 2 D82240 conserved hypothet
888 31 38.3 190 2 A87263 hypothetical prote
889 31 38.3 195 2 F69283 alkyl hydroperoxid
890 31 38.3 197 2 G82973 transcription regu
891 31 38.3 202 2 T36918 probable transcrip
892 31 38.3 202 2 S56267 probable membrane
893 31 38.3 203 2 F91006 hypothetical prote
894 31 38.3 209 2 S51098 probable alkyl hyd
895 31 38.3 209 2 JX0266 platelet aggregati
896 31 38.3 211 2 C81001 probable pyrazinam
897 31 38.3 211 2 H82018 probable nicotinam
898 31 38.3 212 2 B83283 probable transcrip
899 31 38.3 212 2 S37973 DNA-directed RNA p
900 31 38.3 214 2 S59149 H+-transporting tw
901 31 38.3 216 2 JX0265 platelet aggregati
902 31 38.3 217 2 H95995 probable isochoris
903 31 38.3 218 2 JC7220 nuclear prote in SR
904 31 38.3 224 2 T40315 hypothetical prote
905 31 38.3 226 2 B69079 alkyl hydroperoxid

31 38.3 229 2 H70660 probable serine ac
31 38.3 229 2 C87047 conserved hypothet
31 38.3 229 2 JC7219 nuclear prote in SR
31 38.3 231 2 S76456 hypothetical prote
31 38.3 234 2 B72254 glycerol uptake fa
31 38.3 237 2 F70606 probable pkmM prot
31 38.3 246 2 AF0429 conserved hypothet
31 38.3 247 2 T10524 tonoplast intrinsi
31 38.3 249 2 JQ2287 SPCp1 prote in - so
31 38.3 249 2 T09297 tonoplast intrinsi
31 38.3 249 2 D83215 probable ATP-bindi
31 38.3 249 2 H71131 hypothetical prote
31 38.3 251 2 S22202 tonoplast intrinsi
31 38.3 252 2 F83443 hypothetical prote
31 38.3 252 2 A71691 hypothetical prote
31 38.3 253 2 S13718 probable membrane
31 38.3 253 2 T51819 gamma tonoplast in
31 38.3 253 2 D87392 nucleotidyltransfe
31 38.3 254 2 AB0655 phosphatidylglycer
31 38.3 254 2 B72549 hypothetical prote
31 38.3 255 2 JQ2288 SPCp2 prote in - so
31 38.3 255 2 T04389 NBS-LRR type resis
31 38.3 255 2 H97757 hydrolase homolog
31 38.3 256 2 A13589 pca regulon regula
31 38.3 256 2 D86544 NADH (ubiquinone)
31 38.3 256 2 H72078 probable sodium-tr
31 38.3 256 2 A96027 probable ABC trans
31 38.3 260 2 B83833 hypothetical prote
31 38.3 261 2 C86702 prolipoprote in dia
31 38.3 261 2 JC5806 aquaporin 8 - mous
31 38.3 261 2 AF2201 bacteriorhodopsin
31 38.3 262 2 I40221 divIB prote in - Ba
31 38.3 265 2 B70904 hypothetical prote
31 38.3 265 2 AH0755 conserved hypothet
31 38.3 266 2 D82581 probable signal pe
31 38.3 266 2 A29310 MHC class II histo
31 38.3 267 2 H75429 transcription regu
31 38.3 269 2 B32410 mastocytoma protei
31 38.3 270 2 G97299 pyrroline-5-carbox
31 38.3 270 2 T46856 indole-3-glycerol-
31 38.3 272 2 T35231 hypothetical prote
31 38.3 273 2 T11706 NADH2 dehydrogenas
31 38.3 274 2 T11727 NADH2 dehydrogenas
31 38.3 274 2 T25783 hypothetical prote
31 38.3 275 2 D83841 purine nucleoside
31 38.3 276 2 D56406 casein kinase I (E
31 38.3 277 2 S57381 protein disulfide
31 38.3 278 2 A71498 hypothetical prote
31 38.3 278 2 B83354 probable ATP-bindi
31 38.3 279 2 T25801 hypothetical prote
31 38.3 280 2 E87429 conserved hypothet
31 38.3 284 2 S27931 Env/v-mpl fusion p
31 38.3 286 2 B83343 hypothetical prote
31 38.3 287 2 A54026 alcohol sulfotrans
31 38.3 289 2 C83207 probable hydrolase
31 38.3 289 2 F69313 conserved hypothet
31 38.3 291 2 H81729 succinyl-CoA synth
31 38.3 292 2 H75055 hypothetical prote
31 38.3 295 2 T43930 hypothetical prote
31 38.3 296 2 S01441 class II histocomp
31 38.3 297 2 AH0341 probable aldo/keto
31 38.3 299 2 G70887 fbpC1 prote in prec
31 38.3 303 2 AG0603 hypothetical ABC t
31 38.3 303 2 H64820 probable oligopept
31 38.3 303 2 G90742 probable transport
31 38.3 303 2 B85593 probable transport
31 38.3 304 2 B69474 endoglucanase (cel
31 38.3 304 2 A87258 hypothetical prote
31 38.3 307 2 AB3614 glycosyl transfera
31 38.3 309 2 T25875 hypothetical prote
31 38.3 310 2 G81381 probable integral
31 38.3 320 2 S17177 probable G prote in
31 38.3 320 2 A46152 A3 adenosine recep

979 31 38.3 323 2 E71134 hypothetical prote
980 31 38.3 324 2 T07832 probable steroid s
981 31 38.3 328 2 S64098 nucleotide-binding
982 31 38.3 328 2 T37734 SUR4 family protei
983 31 38.3 329 2 S74595 hypothetical prote
984 31 38.3 331 2 B48445 glyceraldehyde-3-p
985 31 38.3 331 2 A71870 integrase/recombin
986 31 38.3 331 2 JT0589 pectinesterase (EC
987 31 38.3 331 2 T25785 hypothetical prote
988 31 38.3 332 2 F75473 conserved hypothet
989 31 38.3 333 2 I65989 G protein-coupled
990 31 38.3 333 2 S12392 fanF protein - Esc
991 31 38.3 333 2 F90172 hypothetical prote
992 31 38.3 334 2 AD0775 probable hydrolase
993 31 38.3 335 2 T25498 hypothetical prote
994 31 38.3 337 2 S68678 adenosine receptor
995 31 38.3 340 2 T13781 NADH2 dehydrogenas
996 31 38.3 342 2 T29245 hypothetical prote
997 31 38.3 343 2 AG2465 hypothetical prote
998 31 38.3 344 2 D97899 hypothetical prote
999 31 38.3 345 1 JH0185 D-amino-acid oxida
1000 31 38.3 345 2 E69678 involved in polyke

ALIGNMENTS

RESULT 1
T30343
irp3 protein - Yersinia enterocolitica
C;Species: Yersinia enterocolitica
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30343
R;Pelludat, C.; Rakin, A.; Jacobi, C.; Schubert, S.; Heesemann, J.
J. Bacteriol. 180, 538-546, 1998
A;Title: The yersiniabactin biosynthetic gene cluster of Yersinia enterocolitica: organi
A;Reference number: Z20833; MUID:98117033; PMID:9457855
A;Accession: T30343
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-365 <PEL>
A;Cross-references: UNIPROT:O54512; UNIPARC:UPI00000B099E; EMBL:Y12527; PIDN:CAA73128.1
C;Genetics:
A;Gene: irp3

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWL 12
||| :|||
Db 290 CETVGPEGVSWL 301

RESULT 2
T17441
hypothetical protein ybtU [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17441; T47049
R;Fetherston, J.D.; Bertolino, V.J.; Perry, R.D.
Mol. Microbiol. 32, 289-299, 1999
A;Title: YbtP and YbtQ: two ABC transporters required for iron uptake in Yersinia pestis
A;Reference number: Z18782; MUID:99248409; PMID:10231486
A;Accession: T17441
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-365 <FET>
A;Cross-references: UNIPROT:Q9Z3C6; UNIPARC:UPI00000B43BB; EMBL:AF091251; NID:g3818595;
R;Buchrieser, C.; Rusnlok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carniel
submitted to the EMBL Data Library, October 1998
A;Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A;Reference number: Z24348

A;Accession: T47049
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-365 <BOC>
A;Cross-references: UNIPARC:UPI00000B43BB; EMBL:AL031866; PIDN:CAA21392.1
A;Experimental source: strain 6/69
C;Genetics:
A;Gene: ybtU

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWL 12
||| :|||
Db 290 CETVGPEGVSWL 301

RESULT 3
AI0232
yersiniabactin biosynthetic protein YbtU [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AI0232
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0232
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-366 <KUR>
A;Cross-references: UNIPROT:Q8ZF25; UNIPARC:UPI00000CD874; GB:AL590842; PIDN:CAC90725.1;
C;Genetics:
A;Gene: ybtU

Query Match 56.8%; Score 46; DB 2; Length 366;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWL 12
||| :|||
Db 291 CETVGPEGVSWL 302

RESULT 4
ABI885
hypothetical protein alr0627 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: ABI885
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB1885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <KUR>
A;Cross-references: UNIPROT:Q8YZ60; UNIPARC:UPI00000CDDEB; GB:BA000019; PIDN:BAB72585.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0627

Query Match 55.6%; Score 45; DB 2; Length 907;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VGPSLMSWL 12


```

QY      3  LVGPSLSMWL 12
      | | | | | : | | :
Db      257  LAGPSLVSMV 266

RESULT 10
G88065
protein T16A1.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G88065
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G88065
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-786 <STO>
A;Cross-references: UNIPROT:P91451; UNIPARC:UPI000007BD2C; GB:chr_II; PIDN:AB37879.1; PI
C;Genetics:
A;Gene: T16A1.2
A;Map position: 2

```

Query Match	51.9%;	Score 42;	DB 2;	Length 786;
Best Local Similarity	70.0%;	Pred. No. 48;		
Matches	7;	Conservative	1;	Mismatches
			2;	Indels
				Gaps
				0;

QY	5	GPSLMSW	LTC	14
			:	
Db	466	GKSLLSW	LKC	475

RESULT 11

S49844

probable membrane protein YDR089w - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein D4495; hypothetical protein YD6652.01; hypothetical
C;Species: *Saccharomyces cerevisiae*
C;Date: 13-Jan-1995 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C;Accession: S49844; S48778; S55838; S67906; S58089
R;Richards, C.; Harris, D.E.
submitted to the EMBL Data Library, November 1994
A;Reference number: S49823
A;Accession: S49844
A;Molecule type: DNA
A;Residues: 1-260 <RIC>
A;Cross-references: UNIPROT:P38966; UNIPARC:UPI0000168A94; EMBL:Z46796; NID:G577794; PID:
R;Coster, F.; Jonniaux, J.L.; Goffeau, A.
submitted to the EMBL Data Library, October 1994
A;Reference number: S48758
A;Accession: S48778
A;Molecule type: DNA
A;Residues: 1-161 <COS>
A;Cross-references: UNIPARC:UPI0000168B72; EMBL:X82086; NID:G558241; PID:G558262
R;Coster, F.; Jonniaux, J.L.; Goffeau, A.
Yeast 11, 673-679, 1995
A;Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr
A;Reference number: S55819; MUID:96093910; PMID:7483840
A;Accession: S55838
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-161 <COW>
A;Cross-references: UNIPARC:UPI0000168B72; EMBL:X82086; NID:G558241; PIDN:CAA57618.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
R;Foury, F.; Jonniaux, J.L.; Purnelle, B.; Coster, F.; Goffeau, A.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67889
A;Accession: S67906
A;Molecule type: DNA
A;Residues: 1-161 <FOU>

A;Cross-references: UNIPARC:UPI0000168B72; EMBL:Z74385; MIPS:YDR089w
A;Experimental source: strain S288C
R;Oliver, K.; Harris, D.
submitted to the EMBL Data Library, July 1995
A;Reference number: S58089
A;Accession: S58089
A;Molecule type: DNA
A;Residues: 249-869 <OLI>
A;Cross-references: UNIPARC:UPI000006B647; EMBL:Z50111; NID:G914872; PID:G914873
A;Experimental source: strain AB972
C;Genetics:
A;Cross-references: SGD:S0002496
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YDR089w
C;Keywords: transmembrane protein
F;777-793/Domain: transmembrane #status predicted <TM1>
F;811-827/Domain: transmembrane #status predicted <TM2>
F;828-844/Domain: transmembrane #status predicted <TM3>

Query Match	51.9%;	Score 42;	DB 2;	Length 869;
Best Local Similarity	42.9%;	Pred. No. 53;		
Matches	6;	Conservative	2;	Mismatches 6; Indels 0; Gaps 0;

Qy 1 CELVGP S LMSW L T C 14
| | : | : |
Db 537 CSLIP PDL T T W K I C 550

Qy 1 CELVGPSLMSWLT 14
| | : | | : | |
Db 537 CSLIPPDLTWKIC 550

RESULT 12
EOBOA

enkephalin e-12 precursor - bovine
N:Alternate names: proenkephalin A precursor
N:Contains: Leu-enkephalin; Met-enkephalin; Met-enkephalin-Arg-Gly-Leu; Met-enkephalin-Arg-Phe
C:Species: Bos primigenius taurus (cattle)
C:Date: 18-Aug-1982 #sequence revision 18-Aug-1982 #text change 09-Jul-2004
C:Accession: A93272; A93273; A33304; A19448; S62377; S13473; S13474; I45950; A01475
R:Noda, M.; Furutani, Y.; Takahashi, H.; Toyosato, M.; Hirose, T.; Inayama, S.; Nakanishi, T.
Nature 295, 202-206, 1982
A:Title: Cloning and sequence analysis of cDNA for bovine adrenal preproenkephalin.
A:Reference number: A93272; MUID:82125409; PMID:6276759
A:Accession: A93272
A:Molecule type: mRNA
A:Residues: 1-263 <NOD>
A:Cross-references: UNIPROT:P01211; UNIPARC:UPI00001315A6; GB:J00005; GB:V00109; NID:g167
A:Experimental source: adrenal medulla
R:Gubler, U.; Seeburg, P.; Hoffman, B.J.; Gage, L.P.; Udenfriend, S.
Nature 295, 206-208, 1982
A:Title: Molecular cloning establishes proenkephalin as precursor of enkephalin-containing peptides.
A:Reference number: A93273; MUID:82125410; PMID:6173760
A:Accession: A93273
A:Molecule type: mRNA
A:Residues: 25-263 <GUB>
A:Cross-references: UNIPARC:UPI0000173414
A:Experimental source: adrenal medulla
R:Watkinson, A.; Young, J.; Varro, A.; Dockray, G.J.
J. Biol. Chem. 264, 3061-3065, 1989
A:Title: The isolation and chemical characterization of phosphorylated enkephalin-containing peptides.
A:Reference number: A33304; MUID:89123418; PMID:2914943
A:Accession: A33304
A:Molecule type: protein
A:Residues: 233-251 <WAT>
A:Cross-references: UNIPARC:UPI0000173415
A:Experimental source: adrenal medulla
R:Jones, B.N.; Shively, J.E.; Kilpatrick, D.L.; Stern, A.S.; Lewis, R.V.; Kojima, K.; Udenfriend, S.
Proc. Natl. Acad. Sci. U.S.A. 79, 2096-2100, 1982
A:Title: Adrenal opioid proteins of 8600 and 12,600 daltons: intermediates in proenkephalin processing.
A:Reference number: A19448; MUID:82197644; PMID:6952256
A:Accession: A19448
A:Molecule type: protein
A:Residues: 25-60;61-62, 'Z', 64-66, 'Z', 68-69, 'Z', 71-73;74, 'Z', 76-78, 'B', 80-86, 'ZZZ', 90-94
A:Cross-references: UNIPARC:UPI0000173416; UNIPARC:UPI0000173417; UNIPARC:UPI0000173418; UNIPARC:UPI0000173419
R:Goumon, Y.; Strub, J.M.; Moniatte, M.; Nullans, G.; Poteur, L.; Hubert, P.; van Dorsselaere, J.
Eur. J. Biochem. 235, 516-525, 1996

A;Title: The C-terminal bisphosphorylated proenkephalin-A- (209-237) -peptide from adrenal
A;Reference number: S62377; MUID:96184524; PMID:8654396
A;Accession: S62377
A;Molecule type: protein
A;Residues: 233-261 <GOU>
A;Cross-references: UNIPARC:UPI000017341C
R;Mindroiu, T.; Carretero, O.A.; Walz, D.; Scicli, A.G.
Biochim. Biophys. Acta 1076, 9-14, 1991
A;Title: Tissue kallikrein processes small proenkephalin peptides.
A;Reference number: S13473; MUID:91098273; PMID:1986798
A;Accession: S13473
A;Molecule type: protein
A;Residues: 104-137 <MIN>
A;Cross-references: UNIPARC:UPI00000352EE
A;Accession: S13474
A;Molecule type: protein
A;Residues: 206-227 <MIW>
A;Cross-references: UNIPARC:UPI00000352EE
R;Comb, M.; Herbert, E.; Crea, R.
Proc. Natl. Acad. Sci. U.S.A. 79, 360-364, 1982
A;Title: partial characterization of the mrna that codes for enkephalins in bovine adren
A;Reference number: I45950; MUID:82197495; PMID:6952189
A;Accession: I45950
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 211-229 <COM>
A;Cross-references: UNIPARC:UPI000016C34D; GB:J00012; NID:gl63435; PIDN:AAA30673.1; PID:
C;Superfamily: proenkephalin
C;Keywords: neuropeptide; opioid peptide; phosphoprotein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;97-101/Product: Met-enkephalin #status experimental <ME1>
F;104-108/Product: Met-enkephalin #status experimental <ME2>
F;133-137/Product: Met-enkephalin #status experimental <ME3>
F;182-189/Product: Met-enkephalin-Arg-Gly-Leu #status predicted <ME1>
F;206-210/Product: Met-enkephalin #status predicted <ME4>
F;226-230/Product: Leu-enkephalin #status predicted <ME1>
F;257-263/Product: Met-enkephalin-Arg-Phe #status predicted <MEF>
F;239,245,247/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 50.6%; Score 41; DB 1; Length 263;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLTLC 14
||| ||| :|||
Db 52 CEGKPLSLKTWETC 65

RESULT 13
A87380
conserved hypothetical protein CC1053 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C;Accession: A87380
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87380
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <STO>
A;Cross-references: UNIPROT:Q9A9D5; UNIPARC:UPI00000C725F; GB:AE005673; NID:gl3422351; F
C;Genetics:
A;Gene: CC1053
C;Superfamily: Alpha/beta hydrolase

Query Match 50.6%; Score 41; DB 2; Length 290;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTLC 14
||| ||| :|||
Db 98 LVGSSMGWLSLSC 109

RESULT 14
H71631
queuine tRNA-ribosyltransferase (tgt) RP721 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: H71631
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Almark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: H71631
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <AND>
A;Cross-references: UNIPROT:Q9ZCK8; UNIPARC:UPI0000136CF9; GB:AJ235273; GB:AJ235269; NID
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: tgt; RP721
C;Superfamily: queuine tRNA-ribosyltransferase

Query Match 50.6%; Score 41; DB 2; Length 361;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ELVGPSLMSW 11
|:| | | | |
Db 325 EILGPMLMTW 334

RESULT 15
S26414
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004
C;Accession: S26414
R;Buffard, D.; Breda, C.; El Turk, J.; Sallaud, C.; Kondorosi, A.; Esnault, R.
submitted to the EMBL Data Library, August 1992
A;Description: Molecular cloning of two chalcone synthase cDNA from alfalfa.
A;Reference number: S26414
A;Accession: S26414
A;Molecule type: mRNA
A;Residues: 1-389 <BUF>
A;Cross-references: UNIPROT:P51078; UNIPARC:UPI0000127922; EMBL:X68106; NID:gl9590; PIDN
C;Superfamily: type III polyketide synthase
C;Keywords: acyltransferase; coenzyme A

Query Match 50.6%; Score 41; DB 2; Length 389;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CELVGPSLMSW 11
||| :||| :|
Db 84 CEIMAPSLDAW 94

RESULT 16
T03888
Na+/Ca2+,K+-exchanging protein homolog C13D9.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03888
R;Nelson, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid C13D9.
A;Reference number: Z15128
A;Accession: T03888
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-703 <NEL>	
A;Cross-references: UNIPROT:O16241; UNIPARC:UPI000007B85B; EMBL:AF016420; NID:g2291168;	
C;Genetics:	
A;Map position: V	
A;Introns: 68/1; 101/3; 145/3; 205/3; 240/1; 368/1; 460/3; 500/1; 621/1	
A;Note: C13D9.7	
Query Match 50.6%; Score 41; DB 2; Length 703;	
Best Local Similarity 53.8%; Pred. No. 63;	
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
Qy 2 ELVGPSLMSWLT 14	
Db 579 ELLGLTVMSWSDC 591	
RESULT 17	
T34414	
hypothetical protein F07E5.8 - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004	
C;Accession: T34414	
R;Du, Z.; Goela, D.	
submitted to the EMBL Data Library, December 1996	
A;Description: The sequence of C. elegans cosmid F07E5.	
A;Reference number: Z21520	
A;Accession: T34414	
A;Status: preliminary; translated from GB/EMBL/DBDJ	
A;Molecule type: DNA	
A;Residues: 1-760 <DUZ>	
A;Cross-references: UNIPROT:P91227; UNIPARC:UPI000007B004; EMBL:U80837; PIDN:AAB37907.1;	
A;Experimental source: strain Bristol N2; clone F07E5	
C;Genetics:	
A;Gene: CESP:F07E5.8	
A;Map position: 2	
A;Introns: 7/1; 720/2	
Query Match 50.6%; Score 41; DB 2; Length 760;	
Best Local Similarity 70.0%; Pred. No. 68;	
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
Qy 5 GPSLMSWLT 14	
Db 440 GRLLSWLKC 449	
RESULT 18	
AB2301	
hypothetical protein asr3961 [imported] - Nostoc sp. (strain PCC 7120)	
C;Species: Nostoc sp. PCC 7120	
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120	
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004	
C;Accession: AB2301	
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi	
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S	
DNA Res. 8, 205-213, 2001	
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana	
A;Reference number: AB1807; MUID:21595285; PMID:11759840	
A;Accession: AB2301	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-84 <KUR>	
A;Cross-references: UNIPROT:Q8YQ74; UNIPARC:UPI00000CE97D; GB:BA000019; PIDN:BAB75660.1;	
A;Experimental source: strain PCC 7120	
C;Genetics:	
A;Gene: asr3961	
Query Match 49.4%; Score 40; DB 2; Length 84;	
Best Local Similarity 50.0%; Pred. No. 12;	
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;	
Qy 1 CELVGPSLMSWLT 14	

Db 64 CRLGTPPQILWLTC 77	
RESULT 19	
T02735	
hypothetical protein At2g29050 [imported] - Arabidopsis thaliana	
N;Alternate names: hypothetical protein T9I4.13	
C;Species: Arabidopsis thaliana (mouse-ear cress)	
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Oct-2004	
C;Accession: T02735; H84691	
R;Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Ron	
submitted to the EMBL Data Library, August 1998	
A;Description: Arabidopsis thaliana chromosome II BAC T9I4 genomic sequence.	
A;Reference number: Z14710	
A;Accession: T02735	
A;Status: translated from GB/EMBL/DBDJ	
A;Molecule type: DNA	
A;Residues: 1-372 <ROU>	
A;Cross-references: UNIPROT:O81073; UNIPARC:UPI00000AB3DD; EMBL:AC005315; NID:g3461834; I	
A;Experimental source: cultivar Columbia	
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I	
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.	
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J	
Nature 402, 761-768, 1999	
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.	
A;Reference number: A84420; MUID:20083487; PMID:10617197	
A;Accession: H84691	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-372 <STO>	
A;Cross-references: UNIPARC:UPI00000AB3DD; GB:AE002093; NID:g3461845; PIDN:AAC33231.1; G	
C;Genetics:	
A;Gene: T9I4.13; At2g29050	
A;Map position: 2	
A;Introns: 137/2; 192/1	
C;Superfamily: AARA protein	
Query Match 49.4%; Score 40; DB 2; Length 372;	
Best Local Similarity 54.5%; Pred. No. 50;	
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;	
Qy 3 LVGPSSLMSWLT 13	
Db 104 LLGPSSLTWVT 114	
RESULT 20	
DEPSXA	
3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) alpha chain - Pseudomonas	
N;Alternate names: 2-oxoisovalerate dehydrogenase (lipoamide) E1-alpha chain; branched-ci	
C;Species: Pseudomonas putida	
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004	
C;Accession: S01317; B36133; S63475	
R;Burns, G.; Brown, T.; Hatter, K.; Idriss, J.M.; Sokatch, J.R.	
Eur. J. Biochem. 176, 311-317, 1988	
A;Title: Similarity of the E1 subunits of branched-chain-oxoacid dehydrogenase from Pseu	
A;Reference number: S01317; MUID:88329084; PMID:3416875	
A;Accession: S01317	
A;Molecule type: DNA	
A;Residues: 1-410 <BUR>	
A;Cross-references: UNIPROT:P09060; UNIPARC:UPI0000172024; EMBL:X13004	
R;Madhusudhan, K.T.; Huang, G.; Burns, G.; Sokatch, J.R.	
J. Bacteriol. 172, 5655-5663, 1990	
A;Title: Transcriptional analysis of the promoter region of the Pseudomonas putida branch	
A;Reference number: A36133; MUID:91008935; PMID:2211503	
A;Accession: B36133	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-17 <MAD>	
A;Cross-references: UNIPARC:UPI0000172025; GB:M33715	
R;Hester, K.; Luo, J.; Burns, G.; Braswell, E.H.; Sokatch, J.R.	
Eur. J. Biochem. 233, 828-836, 1995	
A;Title: Purification of active E1-alpha(2)-beta(2) of Pseudomonas putida branched-chain.	

A;Reference number: S63475; MUID:96085147; PMID:8521848
A;Accession: S63475
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <HES>
A;Cross-references: UNIPARC:UPI0000172026
C;Genetics:
A;Gene: bkdA1
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin
C;Keywords: lipoamide; oxidoreductase; phosphoprotein; thiamin pyrophosphate
F;2-410/Product: 3-methyl-2-oxobutanoate dehydrogenase (lipoamide) alpha chain #status p
F;202-251/Domain: thiamin pyrophosphate-binding domain homology <TPB>
F;313/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 49.4%; Score 40; DB 1; Length 410;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
:||||:|:
Db 297 LGPSLIEWVT 306

RESULT 21
C83365
2-oxoisovalerate dehydrogenase (alpha subunit) PA2247 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: C83365
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: C83365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-410 <STO>
A;Cross-references: UNIPROT:Q9IIM2; UNIPARC:UPI00000D74B0; GB:AE004650; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: bkdA1; PA2247
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin

Query Match 49.4%; Score 40; DB 2; Length 410;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
:||||:|:
Db 297 LGPSLIEWVT 306

RESULT 22
T31134
naphthalene dioxygenase (EC 1.14.12.-) large chain - Sphingomonas aromaticivorans plasm
C;Species: Sphingomonas aromaticivorans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T31134
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati
A;Reference number: Z20992
A;Accession: T31134
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-459 <ROM>
A;Cross-references: UNIPROT:O85843; UNIPARC:UPI000005C95F; EMBL:AF079317; NID:g3378261;
C;Genetics:
A;Gene: bphA1f
A;Genome: plasmid pNLI1
C;Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe

C;Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein
F;80,82,100,103/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 49.4%; Score 40; DB 2; Length 459;
Best Local Similarity 61.5%; Pred. No. 62;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLT 14
||:|||||
Db 180 ELLGPPLKSTLAC 192

RESULT 23
T03889
Na+/Ca2+,K+-exchanging protein homolog C13D9.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03889
R;Nelson, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid C13D9.
A;Reference number: Z15128
A;Accession: T03889
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-651 <NEL>
A;Cross-references: UNIPROT:O16242; UNIPARC:UPI0000079F5D; EMBL:AF016420; NID:g2291168;
C;Genetics:
A;Map position: V
A;Introns: 17/1; 50/3; 94/3; 154/3; 189/1; 316/1; 408/3; 448/1; 569/1
A;Note: C13D9.8

Query Match 49.4%; Score 40; DB 2; Length 651;
Best Local Similarity 46.2%; Pred. No. 86;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLT 14
||:|||||
Db 527 ELLGLTIMAWSNC 539

RESULT 24
T04112
pol protein homolog - maize retrotransposon Opie-2
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04112
R;SanMiguel, P.J.; Tiknonov, A.; Jin, Y.K.; Motchoulskaia, N.; Zakharov, D.; Melake-Berl
Science 274, 765-768, 1996
A;Title: Nested retrotransposons in the intergenic regions of the maize genome.
A;Reference number: Z15219; MUID:97020115; PMID:8864112
A;Accession: T04112
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1068 <SAN>
A;Cross-references: UNIPROT:O24587; UNIPARC:UPI00000A91A6; EMBL:U68408; NID:g1657766; P
C;Genetics:
A;Gene: pol
A;Mobile element: retrotransposon Opie-2
C;Superfamily: retrovirus-related polyprotein

Query Match 49.4%; Score 40; DB 2; Length 1068;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CELVGPSLMSW 11
|:|:|||||
Db 933 CQFLGRSLVSW 943

RESULT 25
S54440
hemin-specific ATP-binding protein [validated] - Yersinia enterocolitica

C;Species: Yersinia enterocolitica
C;Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 05-Oct-2004
C;Accession: S54440
R;Stojiljkovic, I.; Hantke, K.
Mol. Microbiol. 13, 719-732, 1994
A;Title: Transport of haemin across the cytoplasmic membrane through a haemin-specific P
A;Reference number: S54436; MUID:95089707; PMID:7997183
A;Accession: S54440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <STO>
A;Cross-references: UNIPROT:P74981; UNIPARC:UPI00000B721D; EMBL:X77867; NID:g1619622; PI
A;Note: the authors translated the initiation codon GTG for residue 1 as Val
C;Genetics:
A;Gene: hemV
A;Start codon: GTG
C;Function:
A;Description: component of the periplasmic binding-protein-dependent transport system
C;Keywords: ATP; iron transport; nucleotide binding; P-loop
F;27-224/Domain: ATP-binding cassette homology <ABC>
F;44-51/Region: nucleotide-binding motif A (P-loop)

Query Match 48.1%; Score 39; DB 2; Length 266;
Best Local Similarity 54.5%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CELVGPSPMSW 11
|:|:|:|:
Db 68 CQLLGKLNLSW 78

RESULT 26
C85224
Probable LTR retrotransposon (partial) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: C85224
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: C85224
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-306 <STO>
A;Cross-references: UNIPROT:O81859; UNIPARC:UPI000009EF90; GB:NC_001268; NID:g7268774; F
C;Genetics:
A;Gene: AT4g19780
A;Map position: 4
C;Superfamily: retrovirus-related polyprotein

Query Match 48.1%; Score 39; DB 2; Length 306;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPSPMSW 11
|:|:|:|:
Db 168 CMFIGDSLVS 178

RESULT 27
T04759
hypothetical protein T16H5.140 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04759
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04759
A;Molecule type: DNA
A;Residues: 1-306 <BEV>
A;Cross-references: UNIPROT:O81859; UNIPARC:UPI000009EF90; EMBL:AL024486

A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Note: T16H5.140
C;Superfamily: retrovirus-related polyprotein

Query Match 48.1%; Score 39; DB 2; Length 306;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPSPMSW 11
|:|:|:|:
Db 168 CMFIGDSLVS 178

RESULT 28
T48374
UDPG glucosyltransferase-like protein - Arabidopsis thaliana
N;Alternate names: protein F12E4.260
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48374
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24492
A;Accession: T48374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <BEV>
A;Cross-references: UNIPROT:Q9LZD8; UNIPARC:UPI00000A86B9; EMBL:AL162751
A;Experimental source: cultivar Columbia; BAC clone F12E4
C;Genetics:
A;Map position: 5
A;Note: F12E4.260
C;Superfamily: flavonol O3-glucosyltransferase

Query Match 48.1%; Score 39; DB 2; Length 465;
Best Local Similarity 77.8%; Pred. No. 92;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWL 12
|:|:|:|:
Db 268 VDPSLLSWL 276

RESULT 29
S46830
urea transport protein - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHL016c; urea active transporter; urea permease
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1994 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S46830; A40604
R;Favella, T.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 94333.
A;Reference number: S46796
A;Accession: S46830
A;Molecule type: DNA
A;Residues: 1-735 <FAV>
A;Cross-references: UNIPROT:P33413; UNIPARC:UPI000012995C; EMBL:U11582; NID:g2289793; PI
R;ElBerry, H.M.; Majumdar, M.L.; Cunningham, T.S.; Sumrada, R.A.; Cooper, T.G.
J. Bacteriol. 175, 4688-4698, 1993
A;Title: Regulation of the urea active transporter gene (DUR3) in Saccharomyces cerevisiae
A;Reference number: A40604; MUID:93328673; PMID:8335627
A;Accession: A40604
A;Molecule type: DNA
A;Residues: 1-57, 'RLS', 58, 'CS', 64-735 <ELB>
A;Cross-references: UNIPARC:UPI0000168BF4; EMBL:L19875; NID:g388366; PIDN:AAA34582.1; PI
A;Note: sequence extracted from NCBI backbone (NCBIN:135628, NCBIP:135629)
C;Genetics:
A;Gene: SGD:DUR3
A;Cross-references: SGD:S0001008; MIPS:YHL016c
A;Map position: 8L

C;Function:
A;Description: urea transport
C;Keywords: transmembrane protein; urea transport

Query Match 48.1%; Score 39; DB 2; Length 735;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GPSLMSWLTLC 14
|:|||||
Db 467 GLAIMSWLVC 476

RESULT 30
S75152
hypothetical protein slr1767 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75152
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75152
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-112 <KAN>
A;Cross-references: UNIPROT:P73045; UNIPARC:UPI00000C0C7F; EMBL:D90903; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 46.9%; Score 38; DB 2; Length 112;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWLTLC 14
| | | : ||||
Db 64 CRLGIPPQIIWLTC 77

RESULT 31
A46097
GPI-anchor biosynthesis protein FIG-F - human
C;Species: Homo sapiens (man)
C;Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004
C;Accession: A46097
R;Inoue, N.; Kinoshita, T.; Orii, T.; Takeda, J.
J. Biol. Chem. 268, 6882-6885, 1993
A;Title: Cloning of a human gene, FIG-F, a component of glycosyl-phosphatidylinositol an
A;Reference number: A46097; MUID:93216618; PMID:8463218
A;Accession: A46097
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-219 <INO>
A;Cross-references: UNIPROT:Q07326; UNIPARC:UPI0000030AE2; GB:D13435; NID:G303615; PIDN:
C;Genetics:
A;Gene: GDB:PIGF
A;Cross-references: GDB:138466; OMIM:600153
A;Map position: 2p16-2p21

Query Match 46.9%; Score 38; DB 2; Length 219;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
|:|:|:|
Db 130 LLGNLKAWL 139

RESULT 32
C90848

hypothetical protein ECs1755 [imported] - Escherichia coli (strain O157:H7, substrain RIN
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C90848
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C90848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <HAY>
A;Cross-references: UNIPROT:Q8XCB7; UNIPARC:UPI00000D04C2; GB:BA000007; PIDN:BAB35178.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1755
C;Superfamily: yciC protein

Query Match 46.9%; Score 38; DB 2; Length 247;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
| | | : ||||
Db 189 LVAPAVLSWL 198

RESULT 33

B85706
hypothetical protein yciC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85706
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85706
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <STO>
A;Cross-references: UNIPROT:Q8XCB7; UNIPARC:UPI00000D04C2; GB:AE005174; NID:G12514977; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yciC
C;Superfamily: yciC protein

Query Match 46.9%; Score 38; DB 2; Length 247;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
| | | : ||||
Db 189 LVAPAVLSWL 198

RESULT 34

B64873
probable membrane protein yciC - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: B64873; S07798
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cr
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64873
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-247 <BLAT>

A;Cross-references: UNIPROT:P21365; UNIPARC:UPI0000168089; GB:AE000223; GB:U00096; NID:9
A;Experimental source: strain K-12, substrain MGL655
R;Stolitzfus, A.; Leslie, J.F.; Milkman, R.
Genetics 120, 345-358, 1988
A;Title: Molecular evolution of the Escherichia coli chromosome. I. Analysis of structure
A;Reference number: S07794; MUID:89065346; PMID:3058546
A;Accession: S07798
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-99,'RNINYPAGVCGSESQ',115-190,'T',192-201,'P',203-247 <STO>
A;Cross-references: UNIPARC:UPI000013A6BA; EMBL:X13583; NID:G43207; PIDN:CAA31923.1; PID
C;Genetics:
A;Gene: YciC
A;Map position: 28 min
C;Superfamily: YciC protein
C;Keywords: transmembrane protein
F;21-37/Domain: transmembrane #status predicted <TM1>
F;91-107/Domain: transmembrane #status predicted <TM2>
F;124-140/Domain: transmembrane #status predicted <TM3>
F;144-160/Domain: transmembrane #status predicted <TM4>
F;198-214/Domain: transmembrane #status predicted <TM5>
F;229-245/Domain: transmembrane #status predicted <TM6>

Query Match 46.9%; Score 38; DB 2; Length 247;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVGPSSLMSWL 12
Db 189 LVAPAVLSWL 198

RESULT 35
EQHUA
enkephalin precursor - human
N;Alternate names: proenkephalin A precursor
N;Contains: Leu-enkephalin; Met-enkephalin 1; Met-enkephalin 2; Met-enkephalin 3; Met-en
C;Species: Homo sapiens (man)
C;Date: 18-Aug-1982 #sequence revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A93278; A93276; A01474
R;Noda, M.; Teranishi, Y.; Takahashi, H.; Toyosato, M.; Notake, M.; Nakanishi, S.; Numa,
Nature 297, 431-434, 1982
A;Title: Isolation and structural organization of the human preproenkephalin gene.
A;Reference number: A93278; MUID:82195574; PMID:6281660
A;Accession: A93278
A;Molecule type: DNA
A;Residues: 1-267 <NOD>
A;Cross-references: UNIPROT:P01210; UNIPARC:UPI00001315A9; GB:J00123; GB:V00510; NID:G18
A;Note: another intron is present in the 5' end of the noncoding region
R;Comb, M.; Seeburg, P.H.; Adelman, J.; Eiden, L.; Herbert, E.
Nature 295, 663-666, 1982
A;Title: Primary structure of the human Met- and Leu-enkephalin precursor and its mRNA.
A;Reference number: A93276; MUID:82125511; PMID:7057924
A;Accession: A93276
A;Molecule type: mRNA
A;Residues: 1-267 <COM>
A;Cross-references: UNIPARC:UPI00001315A9; GB:J00123; GB:V00510; NID:G182097; PIDN:AAB59
A;Note: this mRNA was obtained from pheochromocytoma, a tumor of the adrenal medulla; t
C;Genetics:
A;Gene: GDB:PENK
A;Cross-references: GDB:120270; OMIM:131330
A;Map position: 8q11.23-8q12
A;Introns: 46/3
C;Superfamily: proenkephalin
C;Keywords: neuropeptide; opioid peptide
F;1-24/Domain: signal sequence #status predicted <SIG>
F;100-104/Product: Met-enkephalin #status predicted <ME1>
F;107-111/Product: Met-enkephalin #status predicted <ME2>
F;136-140/Product: Met-enkephalin #status predicted <ME3>
F;186-193/Product: Met-enkephalin-Arg-Gly-Leu #status predicted <MEL>
F;210-214/Product: Met-enkephalin #status predicted <ME4>
F;230-234/Product: Leu-enkephalin #status predicted <LE1>
F;261-267/Product: Met-enkephalin-Arg-Phe #status predicted <MEF>

Query Match 46.9%; Score 38; DB 1; Length 267;
Best Local Similarity 57.1%; Pred. No. 79;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLT 14
Db 52 CEGKLPSLKIWET 65

RESULT 36
C95847
hypothetical transmembrane protein SMB20040 [imported] - Sinorhizobium meliloti (strain
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95847
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: C95847
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <KUR>
A;Cross-references: UNIPROT:Q92XB9; UNIPARC:UPI00000D4784; GB:AL591985; PIDN:CAC48443.1;
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB20040
A;Genome: plasmid

Query Match 46.9%; Score 38; DB 2; Length 286;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSSLMSWL 12
Db 203 LAGPSAASWL 212

RESULT 37
T09355
hypothetical protein F23K16.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T09355
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16652
A;Accession: T09355
A;Molecule type: DNA
A;Residues: 1-342 <BEV>
A;Cross-references: UNIPROT:Q9SVB8; UNIPARC:UPI000009EFB1; EMBL:AL078620; GSPDB:GN000062;
A;Experimental source: cultivar Columbia; BAC clone F23K16
C;Genetics:
A;Gene: ATSP:F23K16.20
A;Map position: 4
A;Introns: 45/1; 106/3; 173/3; 227/3

Query Match 46.9%; Score 38; DB 2; Length 342;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSSLMSWL 13
Db 111 | | | |

Db 203 LVGPFLDYWLT 213

RESULT 38

T12379

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Panorpa japonica mitochondrion

C;Species: mitochondrion Panorpa japonica

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C;Accession: T12379

R;Ogai, H.

submitted to the EMBL Data Library, March 1998

A;Description: A genetic evidence showing the identity between form japonica and form kl

A;Reference number: Z17500

A;Accession: T12379

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-357 <OGA>

A;Cross-references: UNIPROT:O63702; UNIPARC:UPI0000092D34; EMBL:AF056494; NID:g3063638;

C;Genetics:

A;Genome: mitochondrion

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 46.9%; Score 38; DB 2; Length 357;

Best Local Similarity 70.0%; Pred. No. 1e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVGPSLMSWL 12

Db 213 VLGGSLMSWL 222

RESULT 39

B29336

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Rhodobacter capsulatus

C;Species: Rhodobacter capsulatus

C;Date: 31-Dec-1988 #sequence_revision 22-Jul-1994 #text_change 09-Jul-2004

C;Accession: B29336; B25405; S09373

R;Davidson, E.; Daldal, F.

J. Mol. Biol. 195, 13-24, 1987

A;Title: Primary structure of the bc-1 complex of Rhodopseudomonas capsulata. Nucleotide

A;Reference number: A92938; MUID:88011223; PMID:2821268

A;Accession: B29336

A;Molecule type: DNA

A;Residues: 1-437 <DAV>

A;Cross-references: UNIPROT:P08502; UNIPARC:UPI000016FE4D; EMBL:X05630; NID:g46093; PIDN

R;Gabellini, N.; Sebald, W.

Eur. J. Biochem. 154, 569-579, 1986

A;Title: Nucleotide sequence and transcription of the fbc operon from Rhodopseudomonas s

A;Reference number: A91162; MUID:86136096; PMID:3004982

A;Note: source is designated as Rhodopseudomonas sphaeroides

A;Accession: B25405

A;Molecule type: DNA

A;Residues: 1-66,'ID',69-280,'I',282-437 <GAB>

A;Cross-references: UNIPARC:UPI000016FE34; EMBL:X03476; NID:g46007; PIDN:CAA27195.1; PID

C;Genetics:

A;Gene: fbcB; petB

C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase

F;26-381/Domain: cytochrome b homology <CBH>

F;26-225/Domain: cytochrome b6 homology <CB6>

F;51-67/Domain: transmembrane #status predicted <TM1>

F;96-114/Domain: transmembrane #status predicted <TM2>

F;134-150/Domain: transmembrane #status predicted <TM3>

F;146-193/Domain: periplasmic #status predicted <PER1>

F;195-217/Domain: transmembrane #status predicted <TM4>

F;245-381/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F;253-269/Domain: transmembrane #status predicted <TM5>

F;270-329/Domain: periplasmic #status predicted <PER2>

F;330-346/Domain: transmembrane #status predicted <TM6>

F;365-383/Domain: transmembrane #status predicted <TM7>

F;395-411/Domain: transmembrane #status predicted <TM8>

F;97,198/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F;111,212/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 46.9%; Score 38; DB 1; Length 437;

Best Local Similarity 55.6%; Pred. No. 1.3e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWL 12

Db 172 IGPSIQAWL 180

RESULT 40

B69004

hypothetical protein MTH1028 - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: B69004

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: B69004

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-441 <MTH>

A;Cross-references: UNIPROT:O27107; UNIPARC:UPI0000062B87; GB:AE000875; GB:AE000666; NID

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1028

Query Match 46.9%; Score 38; DB 2; Length 441;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLT 13

Db 194 DVVGPATQSWRT 205

RESULT 41

AG0308

isochorismate synthase (EC 5.4.99.6) [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 12-Jun-2003

C;Accession: AG0308

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AG0308

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-455 <KUR>

A;Cross-references: UNIPARC:UPI00000CD92A; GB:AL590842; PIDN:CAC91331.1; PID:g15980520;

C;Genetics:

A;Gene: menF

C;Superfamily: isochorismate synthase

C;Keywords: intramolecular transferase; isomerase

Query Match 46.9%; Score 38; DB 2; Length 455;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLT 13

Db 35 LVGRQLLEWLT 45

RESULT 42
AI3367
multidrug resistance protein B [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AI3367
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessa Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AI3367
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <KUR>
A;Cross-references: UNIPROT:Q8YH77; UNIPARC:UPI0000057E89; GB:AE008917; PIDN:AAL52108.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0927
A;Map position: I

Query Match 46.9%; Score 38; DB 2; Length 472;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLT 13
Db 113 ILGPTLGGWLT 123

RESULT 43
T48630
high affinity nitrate transporter-like protein - Arabidopsis thaliana
N;Alternate names: protein T15N1.60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48630
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T48630
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <BEV>
A;Cross-references: UNIPROT:Q9LYK2; UNIPARC:UPI0000048B31; EMBL:AL163792
A;Experimental source: cultivar Columbia; BAC clone T15N1
C;Genetics:
A;Map position: 5
A;Introns: 395/2
A;Note: T15N1.60
C;Superfamily: nitrate transporter component

Query Match 46.9%; Score 38; DB 2; Length 493;
Best Local Similarity 41.2%; Pred. No. 1.4e+02;
Matches 7; Conservative 6; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CELVGP----SLMSWLT 13
Db 99 CDLIGPRTSSAILLSFLT 115

RESULT 44
H83254
probable MFS transporter PA3137 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2004
C;Accession: H83254
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: H83254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-499 <STO>
A;Cross-references: UNIPROT:Q9HZ88; UNIPARC:UPI000000C5925; GB:AE004737; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3137

Query Match 46.9%; Score 38; DB 2; Length 499;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VGPSLMSWLT 13
Db 138 IGPTLGGWLT 147

RESULT 45
A70022
multidrug-efflux transporter homolog yusP - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A70022
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A70022
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-541 <KUN>
A;Cross-references: UNIPROT:O32182; UNIPARC:UPI0000060A25; GB:Z99120; GB:AL009126; NID:9
A;Experimental source: strain 168
C;Genetics:
A;Gene: yusP
C;Superfamily: multidrug-efflux transporter
C;Keywords: antibiotic resistance; transmembrane protein

Query Match 46.9%; Score 38; DB 1; Length 541;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWLT 13
Db 43 DLGGLSMMTWLT 54

RESULT 46
T22072
hypothetical protein F41D3.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22072
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19508
A;Accession: T22072
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-594 <WIL>

A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB4; EMBL:Z81537; PIDN:CAB04376.1.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.10
A;Map position: 1
A;Introns: 80/3; 127/1; 194/2; 250/1; 356/3; 399/1; 437/2; 507/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 594;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLT 14
:|:|:|:|:
Db 400 CEVLYPSLVKGMVC 413

RESULT 47

T22067
hypothetical protein F41D3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22067
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19508
A;Accession: T22067
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-644 <WIL>
A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB3; EMBL:Z81537; PIDN:CAB04376.1.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.5
A;Map position: 1
A;Introns: 42/2; 130/3; 177/1; 244/2; 300/1; 406/3; 449/1; 487/2; 557/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 644;
Best Local Similarity 42.9%; Pred. No. 1.8e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 CELVGPSLMSWLT 14
:|:|:|:|:
Db 450 CEVLYPSLVKGMVC 463

RESULT 48

H75485
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75485
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-718 <WHI>
A;Cross-references: UNIPROT:Q9RWG1; UNIPARC:UPI00000C17E3; GB:AE001927; GB:AE000513; NID:
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0707
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0707

Query Match 46.9%; Score 38; DB 2; Length 718;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLT 13
:|:|:|:|:
Db 296 QLLGSGLLEWLT 307

RESULT 49

T04970
probable potassium transport protein KT5 - Arabidopsis thaliana
N;Alternate names: protein T16L1.20
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04970
R;Bevan, M.; Obermaier, B.; Deutschenbaur, S.; Piravandi, E.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, November 1998
A;Reference number: Z15393
A;Accession: T04970
A;Molecule type: DNA
A;Residues: 1-846 <BEV>
A;Cross-references: UNIPROT:Q8LPL8; UNIPARC:UPI000009E991; EMBL:AL031394
A;Experimental source: cultivar Columbia; BAC clone T16L1
C;Genetics:
A;Map position: 4
A;Introns: 98/3; 175/1; 268/1; 355/1; 372/3; 409/1; 494/1; 668/2
A;Note: T16L1.20
C;Superfamily: barley probable potassium transport protein HAK1
C;Keywords: ion transport

Query Match 46.9%; Score 38; DB 2; Length 846;
Best Local Similarity 41.7%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLT 14
:|:|:|:|:
Db 295 VLGPALLWFFC 306

RESULT 50

B96553
hypothetical protein F5D21.7 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B96553
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B96553
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-941 <STO>
A;Cross-references: UNIPROT:Q9C8K0; UNIPARC:UPI00000A5EC3; GB:AE005173; NID:gl0092355;
C;Genetics:
A;Gene: F5D21.7
A;Map position: 1

Query Match 46.9%; Score 38; DB 2; Length 941;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWL 12
:|:|:|:|:
Db 845 KLVGPRDLSQL 855

Job time : 31.0915 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-67

Perfect score: 81

Sequence: 1 CELVGPSLMSWLTC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	50	61.7	508	2	Q6GNG6_XENLA	Q6gng6 xenopus lae
2	50	61.7	526	2	Q4V789_XENTR	Q4v789 xenopus tro
3	49	60.5	199	1	RNS111_HUMAN	Q8taa1 homo sapien
4	49	60.5	199	2	Q5GAN5_HUMAN	Q5gan5 homo sapien
5	49	60.5	500	2	Q9VKC2_DROME	Q9vkc2 drosophila
6	49	60.5	1510	2	Q7XLV1_ORYSA	Q7xly1 oryza sativ
7	47	58.0	167	2	Q7XFQ6_ORYSA	Q7xfq6 oryza sativ
8	47	58.0	167	2	Q8W5F2_ORYSA	Q8w5f2 oryza sativ
9	47	58.0	745	2	Q9HEG1_NEUCR	Q9he61 neurospora
10	46	56.8	365	2	O54512_YEREN	O54512 yersinia en
11	46	56.8	365	2	Q9Z3C6_YERPE	Q9z3c6 yersinia pe
12	46	56.8	366	2	Q8ZF25_YERPE	Q8zf25 yersinia pe
13	46	56.8	386	2	Q8D0C4_YERPE	Q8d0c4 yersinia pe
14	46	56.8	452	2	Q55WU3_CRYNE	Q55wu3 cryptococcu
15	46	56.8	452	2	Q5KJF5_CRYNE	Q5kjf5 cryptococcu
16	46	56.8	716	2	Q8VXB0_ORYSA	Q8vxb0 oryza sativ
17	46	56.8	778	2	Q652J4_ORYSA	Q652j4 oryza sativ
18	46	56.8	1215	2	Q4P7B5_USTMA	Q4p7b5 ustilago ma
19	45	55.6	305	2	Q6A7C7_PROAC	Q6a7c7 propionibac
20	45	55.6	328	2	Q6NEV8_CORDI	Q6nev8 corynebacte
21	45	55.6	497	2	Q8GAG2_ARTNI	Q8gag2 arthrobacte
22	45	55.6	524	1	ARHG3_MOUSE	Q91x46 mus musculu
23	45	55.6	526	1	ARHG3_HUMAN	Q9nr81 homo sapien
24	45	55.6	526	1	ARHG3_MACFA	Q9n0a8 macaca fasc
25	45	55.6	526	2	Q5R6F2_PONPY	Q5r6f2 pongo pygma
26	45	55.6	537	2	Q59F00_HUMAN	Q59f00 homo sapien
27	45	55.6	566	2	Q4QQQ0_HUPIEN	Q4qgq0 homo sapien
28	45	55.6	566	2	Q4QQI5_HUMAN	Q4qqi5 homo sapien
29	45	55.6	734	2	Q52BN4_MAGGR	Q52bn4 magnaporth
30	45	55.6	907	2	Q8YZ60_ANASP	Q8yz60 anabaena sp
31	44	54.3	481	2	Q5TM27_MACMU	Q5tm27 macaca mula

32	44	54.3	481	2	Q8GAI0_ARTNI	Q8gai0 arthrobacte
33	44	54.3	506	2	Q504F6_BRARE	Q504f6 brachydanio
34	44	54.3	836	2	Q8SNA3_MOUSE	Q8sna3 mus musculu
35	44	54.3	1029	2	Q5SQ40_HUMAN	Q5sq40 homo sapien
36	44	54.3	1077	2	Q5SQ42_HUMAN	Q5sq42 homo sapien
37	44	54.3	1078	2	Q5SQ43_HUMAN	Q5sq43 homo sapien
38	44	54.3	1096	2	Q6MG49_RAT	Q6mg49 rattus norv
39	44	54.3	1098	2	Q9WTN8_RAT	Q9wtm8 rattus norv
40	44	54.3	1126	2	Q5HYL9_HUMAN	Q5hyl9 homo sapien
41	44	54.3	1126	2	Q96SA6_HUMAN	Q96sa6 homo sapien
42	44	54.3	1126	2	Q9BCN4_HUMAN	Q9bcn4 homo sapien
43	44	54.3	1132	1	BAT3_HUMAN	P46379 homo sapien
44	44	54.3	1132	2	Q5STC1_HUMAN	Q5stc1 homo sapien
45	44	54.3	1154	2	Q9Z1R2_MOUSE	Q9z1r2 mus musculu
46	44	54.3	1162	2	Q5SQ39_HUMAN	Q5sq39 homo sapien
47	44	54.3	1229	2	Q9S874_HUMAN	Q9s874 homo sapien
48	44	54.3	1229	2	Q5SQ35_HUMAN	Q5sq35 homo sapien
49	43	53.1	108	2	Q4HCM7_9DEIO	Q4hcm7 deinococcus
50	43	53.1	224	2	Q9Z537_STRCO	Q9z537 streptomyce
51	43	53.1	240	2	Q8TQY4_METAC	Q8tqy4 methanosarc
52	43	53.1	388	2	Q6D048_ERWCT	Q6d048 erwinia car
53	43	53.1	457	2	Q6W1T9_RHISN	Q6w1t9 rhizobium s
54	43	53.1	476	2	Q6I3C3_BACAN	Q6i3c3 bacillus an
55	43	53.1	487	2	Q9SK03_MACFA	Q9sk03 macaca fasc
56	43	53.1	505	2	Q73ZV0_MYCPA	Q73zv0 mycobacteri
57	43	53.1	526	2	Q8H0S2_ARATH	Q8h0s2 arabidopsis
58	43	53.1	526	2	Q941F0_ARATH	Q941f0 arabidopsis
59	43	53.1	537	2	Q4MN16_BACCE	Q4mn16 bacillus ce
60	43	53.1	537	2	Q6HNF1_BACHK	Q6hnf1 bacillus th
61	43	53.1	537	2	Q73DI1_BACCI	Q73di1 bacillus ce
62	43	53.1	537	2	Q63FY3_BACCC	Q63fy3 bacillus ce
63	42.5	52.5	262	1	NASD_KLEOX	P39459 klebsiella
64	42.5	52.5	268	2	Q8Y2F7_RALSO	Q8y2f7 ralstonia s
65	42.5	52.5	571	2	Q7SC40_NEUCR	Q7sc40 neurospora
66	42	51.9	74	2	Q9SL31_HORSE	Q9sl31 equus cabal
67	42	51.9	139	2	Q60903_MOUSE	Q60903 mus musculu
68	42	51.9	161	2	Q813S5_BACCR	Q813s5 bacillus ce
69	42	51.9	162	2	Q4MT25_BACCE	Q4mt25 bacillus ce
70	42	51.9	248	1	Y4UH_RHISN	Q53200 rhizobium s
71	42	51.9	271	2	Q94UH0_9HEMI	Q94uh0 trialeurode
72	42	51.9	289	2	Q4J4I1_AZOVI	Q4j4i1 azotobacter
73	42	51.9	366	2	Q8FGE7_ECOL6	Q8fge7 escherichia
74	42	51.9	401	2	O00572_HUMAN	O00572 homo sapien
75	42	51.9	419	1	PELI2_MOUSE	Q8bst6 mus musculu
76	42	51.9	419	2	Q6P4Y5_XENTR	Q6p4y5 xenopus tro
77	42	51.9	420	1	PELI2_HUMAN	Q9hat8 homo sapien
78	42	51.9	428	2	Q5VXH9_HUMAN	Q5vxx9 homo sapien
79	42	51.9	428	2	Q6DHJ8_BRARE	Q6dhj8 brachydanio
80	42	51.9	449	2	Q9FFJ8_ARATH	Q9ffj8 arabidopsis
81	42	51.9	487	1	BAT1_HUMAN	P82251 homo sapien
82	42	51.9	487	2	Q53FY4_HUMAN	Q53fy4 homo sapien
83	42	51.9	497	1	CP26A_HUMAN	O43174 homo sapien
84	42	51.9	497	2	Q5VXI0_HUMAN	Q5vxi0 homo sapien
85	42	51.9	517	2	Q7YSY4_AEDAE	Q7ysy4 aedes aegyp
86	42	51.9	518	2	Q89VW4_BRAJA	Q89vw4 bradyrhizob
87	42	51.9	519	2	Q5DTF7_MOUSE	Q5dtf7 mus musculu
88	42	51.9	520	2	Q6NCF8_RHOPA	Q6ncf8 rhodopseudo
89	42	51.9	539	2	O22252_ARATH	O22252 arabidopsis
90	42	51.9	542	2	Q6APSA4_DESPS	Q6aps4 desulfotale
91	42	51.9	543	2	Q7XM07_ORYSA	Q7xm07 oryza sativ
92	42	51.9	549	2	Q8CIG1_MOUSE	Q8cig1 mus musculu
93	42	51.9	550	2	Q8R1D1_MOUSE	Q8r1d1 mus musculu
94	42	51.9	564	2	Q8BJ29_MOUSE	Q8bj29 mus musculu
95	42	51.9	621	2	Q03192_YEAST	Q03192 saccharomyc
96	42	51.9	632	2	Q580Y2_9TRYP	Q580y2 trypanosoma
97	42	51.9	632	2	Q580Y3_9TRYP	Q580y3 trypanosoma
98	42	51.9	632	2	Q580Y5_9TRYP	Q580y5 trypanosoma
99	42	51.9	686	2	Q07492_9TRYP	Q07492 trypanosoma
100	42	51.9	686	2	Q8WPU7_9TRYP	Q8wpu7 trypanosoma
101	42	51.9	786	2	P91451_CABEL	P91451 caenorhabdi
102	42	51.9	845	2	Q7XLJ6_ORYSA	Q7xlj6 oryza sativ
103	42	51.9	869	1	YD95_YEAST	P38966 saccharomyc
104	42	51.9	1005	2	Q75KS9_ORYSA	Q75ks9 oryza sativ

Q8gai0	arthrobacte
Q504f6	brachydanio
Q8sna3	mus musculu
Q5sq40	homo sapien
Q5sq42	homo sapien
Q5sq43	homo sapien
Q6mg49	rattus norv
Q9wtm8	rattus norv
Q5hyl9	homo sapien
Q96sa6	homo sapien
Q9bcn4	homo sapien
P46379	homo sapien
Q5stc1	homo sapien
Q9z1r2	mus musculu
Q5sq39	homo sapien
Q9s874	homo sapien
Q5sq35	homo sapien
Q4hcm7	deinococcus
Q9z537	streptomyce
Q8tqy4	methanosarc
Q6d048	erwinia car
Q6w1t9	rhizobium s
Q6i3c3	bacillus an
Q9sk03	macaca fasc
Q73zv0	mycobacteri
Q8h0s2	arabidopsis
Q941f0	arabidopsis
Q4mn16	bacillus ce
Q6hnf1	bacillus th
Q73di1	bacillus ce
Q63fy3	bacillus ce
P39459	klebsiella
Q8y2f7	ralstonia s
Q7sc40	neurospora
Q9sl31	equus cabal
Q60903	mus musculu
Q813s5	bacillus ce
Q4mt25	bacillus ce
Q53200	rhizobium s
Q94uh0	trialeurode
Q4j4i1	azotobacter
Q8fge7	escherichia
O00572	homo sapien
Q8bst6	mus musculu
Q6p4y5	xenopus tro
Q9hat8	homo sapien
Q5vxx9	homo sapien
Q6dhj8	brachydanio
Q9ffj8	arabidopsis
P82251	homo sapien
Q53fy4	homo sapien
O43174	homo sapien
Q5vxi0	homo sapien
Q7ysy4	aedes aegyp
Q89vw4	bradyrhizob
Q5dtf7	mus musculu
Q6ncf8	rhodopseudo
O22252	arabidopsis
Q6aps4	desulfotale
Q7xm07	oryza sativ
Q8cig1	mus musculu
Q8r1d1	mus musculu
Q8bj29	mus musculu
Q03192	saccharomyc
Q580y2	trypanosoma
Q580y3	trypanosoma
Q580y5	trypanosoma
Q07492	trypanosoma
Q8wpu7	trypanosoma
P91451	caenorhabdi
Q7xlj6	oryza sativ
P38966	saccharomyc
Q75ks9	oryza sativ

105	42	51.9	1128	2	Q850V9_ORYSA	Q850V9	oryza sativ	178	40	49.4	439	2	Q6AB13_PROAC	Q6ab13	propionibac
106	42	51.9	1299	2	Q8H8K3_ORYSA	Q8h8k3	oryza sativ	179	40	49.4	445	1	PELI3_MOUSE	Q8bxr6	mus musculus
107	42	51.9	1411	2	Q5KQF3_ORYSA	Q5kqf3	oryza sativ	180	40	49.4	447	2	Q9ZSX0_MAIZE	Q9zsx0	zea mays (m
108	42	51.9	1539	2	Q7XP45_ORYSA	Q7xp45	oryza sativ	181	40	49.4	459	2	Q85843_SPHAR	Q85843	sphingomona
109	42	51.9	1577	2	Q60DR2_ORYSA	Q60dr2	oryza sativ	182	40	49.4	461	2	Q53D15_9GAMA	Q53dl5	macaca fusc
110	42	51.9	1584	2	Q8LMY8_ORYSA	Q8lmy8	oryza sativ	183	40	49.4	469	1	PELI3_HUMAN	Q8n2h9	homo sapien
111	42	51.9	1739	2	Q7XXJ0_ORYSA	Q7xxj0	oryza sativ	184	40	49.4	477	1	SYE_SYNPX	Q7u581	synechococc
112	42	51.9	1975	2	Q84MV2_ORYSA	Q84mv2	oryza sativ	185	40	49.4	519	2	Q8W2S4_ORYSA	Q8w2s4	oryza sativ
113	42	51.9	2011	2	Q8H851_ORYSA	Q8h851	oryza sativ	186	40	49.4	524	2	Q51VA5_MAGGR	Q5iva5	magnaporthe
114	41.5	51.2	187	2	Q7UP94_RHOBA	Q7up94	rhodopirell	187	40	49.4	537	2	Q81HX2_BACCR	Q81hx2	bacillus ce
115	41	50.6	121	2	Q562B8_RAT	Q562b8	rattus norv	188	40	49.4	541	2	Q59EH4_HUMAN	Q59eh4	homo sapien
116	41	50.6	126	2	Q5HXZ4_GLUOX	Q5hxz4	gluconobact	189	40	49.4	554	1	ZN426_HUMAN	Q9buy5	homo sapien
117	41	50.6	144	2	Q9GAB0_9BIVA	Q9gab0	pitarr morrh	190	40	49.4	578	2	Q8LMP6_ORYSA	Q8lmp6	oryza sativ
118	41	50.6	144	2	Q62AY7_BURMA	Q62ay7	burkholderi	191	40	49.4	647	2	Q61M45_CABBR	Q61m45	caenorhabdi
119	41	50.6	263	1	PENK_BOVIN	P01211	bos taurus	192	40	49.4	651	2	O16242_CABEL	O16242	caenorhabdi
120	41	50.6	285	2	Q8C9D2_MOUSE	Q8c9d2	mus musculus	193	40	49.4	691	2	Q8VY34_MAIZE	Q8vy34	zea mays (m
121	41	50.6	290	2	Q9A9D5_CAUCR	Q9a9d5	caulobacter	194	40	49.4	713	2	Q59GS5_HUMAN	Q59gs5	homo sapien
122	41	50.6	291	2	Q8CBQ3_MOUSE	Q8cbq3	mus musculus	195	40	49.4	723	2	Q6CSV9_KLJULA	Q6csv9	kluuyveromyc
123	41	50.6	361	1	TGT_RICPR	Q9zck8	rickettsia	196	40	49.4	809	2	Q94I66_ORYSA	Q94i66	oryza sativ
124	41	50.6	361	1	TGT_RICTY	Q68w26	rickettsia	197	40	49.4	852	2	Q8DGP5_SYNEL	Q8dgp5	synechococc
125	41	50.6	389	1	CHS5_MEDSA	P51078	medicago sa	198	40	49.4	877	2	Q8W2U1_ORYSA	Q8w2u1	oryza sativ
126	41	50.6	405	2	Q5ZKT7_CHICK	Q5zkt7	gallus gall	199	40	49.4	1010	2	Q8H8I0_ORYSA	Q8h8i0	oryza sativ
127	41	50.6	418	1	PELI1_HUMAN	Q96fa3	homo sapien	200	40	49.4	1048	2	Q8VY36_MAIZE	Q8vy36	zea mays (m
128	41	50.6	418	1	PELI1_MOUSE	Q8c669	mus musculus	201	40	49.4	1063	2	Q9XEJ4_MAIZE	Q9xej4	zea mays (m
129	41	50.6	418	2	Q53T26_HUMAN	Q53t26	homo sapien	202	40	49.4	1068	2	O24587_MAIZE	O24587	zea mays (m
130	41	50.6	418	2	Q5SRW7_MOUSE	Q5srw7	mus musculus	203	40	49.4	1077	1	ADCY4_HUMAN	Q8nfm4	homo sapien
131	41	50.6	418	2	Q7ZXU3_XENLA	Q7zxu3	xenopus lae	204	40	49.4	1105	2	Q86TZ7_HUMAN	Q86tz7	homo sapien
132	41	50.6	419	2	Q6GQ57_XENLA	Q6gq57	xenopus lae	205	40	49.4	1156	2	Q7XS40_ORYSA	Q7xs40	oryza sativ
133	41	50.6	428	2	Q5Z129_NOCFA	Q5z129	nocardia fa	206	40	49.4	1358	2	Q6L3H0_SOLDE	Q6l3h0	solanum dem
134	41	50.6	440	2	Q4SBZ1_TETNG	Q4sbz1	tetraodon n	207	40	49.4	1481	2	Q4TC89_TETNG	Q4tc89	tetraodon n
135	41	50.6	471	2	Q8ZXX1_PYRAE	Q8zxx1	pyrobaculum	208	40	49.4	1512	2	Q7XBD2_MAIZE	Q7xbd2	zea mays (m
136	41	50.6	487	1	BAT1_MOUSE	Q9qxa6	mus musculus	209	40	49.4	1549	2	Q7XWL7_ORYSA	Q7xwl7	oryza sativ
137	41	50.6	487	1	BAT1_RAT	P82252	rattus norv	210	40	49.4	1553	2	Q8S479_MAIZE	Q8s479	zea mays (m
138	41	50.6	487	2	Q4KM04_RAT	Q4km04	rattus norv	211	40	49.4	1657	2	Q8W0X4_MAIZE	Q8w0x4	zea mays (m
139	41	50.6	494	2	Q4LSG0_9BURK	Q4ls90	burkholderi	212	40	49.4	1760	2	Q7XBC6_MAIZE	Q7xbc6	zea mays (m
140	41	50.6	503	1	K1173_MOUSE	Q9dl93	mus musculus	213	40	49.4	1778	2	Q9AY98_ORYSA	Q9ay98	oryza sativ
141	41	50.6	504	1	K1173_HUMAN	Q9hcb6	homo sapien	214	40	49.4	1866	2	Q7XHI2_ORYSA	Q7xhi2	oryza sativ
142	41	50.6	703	2	O16241_CAEEL	O16241	caenorhabdi	215	40	49.4	1866	2	Q8S5R3_ORYSA	Q8s5r3	oryza sativ
143	41	50.6	747	2	Q8CB19_MOUSE	Q8cb19	mus musculus	216	40	49.4	1892	2	Q8H6I8_MAIZE	Q8h6i8	zea mays (m
144	41	50.6	747	2	Q7TPX6_MOUSE	Q7tpx6	mus musculus	217	40	49.4	1896	2	Q84MR5_ORYSA	Q84mr5	oryza sativ
145	41	50.6	757	2	Q8VXA8_ORYSA	Q8vxa8	oryza sativ	218	40	49.4	2319	2	Q8H6I4_MAIZE	Q8h6i4	zea mays (m
146	41	50.6	760	2	P91227_CAEEL	P91227	caenorhabdi	219	40	49.4	2397	2	Q6Q7Y4_PARPR	Q6q7y4	paramecium
147	41	50.6	867	2	Q7XPL3_ORYSA	Q7xpl3	oryza sativ	220	39.5	48.8	367	2	Q9KW55_PECCE	Q9kw55	pectobacter
148	41	50.6	879	1	MASTL_HUMAN	Q96gx5	homo sapien	221	39.5	48.8	367	2	Q9WXA3_ERWCA	Q9wxa3	erwinia car
149	41	50.6	1209	2	Q7SZF4_BRARE	Q7szf4	brachydanio	222	39	48.1	82	2	Q8L5T4_QUESTU	Q8l5t4	quercus sub
150	40.5	50.0	291	2	Q93JM9_9BURK	Q93jm9	cupriavidus	223	39	48.1	136	2	Q5U6D3_BETVU	Q5u6d3	beta vulgar
151	40	49.4	68	2	Q6EB12_CAMJE	Q6eb12	campylobact	224	39	48.1	155	2	Q9R880_9GAMM	Q9r880	cycloclasti
152	40	49.4	84	2	Q8YQ74_ANASP	Q8yq74	anabaena sp	225	39	48.1	155	2	Q9ZFW4_9GAMM	Q9zfw4	cycloclasti
153	40	49.4	89	2	Q8N6I6_HUMAN	Q8n6i6	homo sapien	226	39	48.1	158	2	Q7QFK5_ANOGA	Q7qfk5	anopheles g
154	40	49.4	104	2	Q5HYL7_HUMAN	Q5hy17	homo sapien	227	39	48.1	160	2	Q5LU30_SILPO	Q5lu30	silicibacte
155	40	49.4	109	2	Q9B034_9VIRU	Q9b034	bacterioph	228	39	48.1	161	2	Q6HKV8_BACHK	Q6hkv8	bacillus th
156	40	49.4	129	2	Q4LBS8_SODGL	Q4lbs8	sodalis glo	229	39	48.1	161	2	Q73AQ6_BACC1	Q73aq6	bacillus ce
157	40	49.4	179	2	Q5RCD5_PONPY	Q5rcd5	pongo pygma	230	39	48.1	161	2	Q63DE6_BACCCZ	Q63de6	bacillus ce
158	40	49.4	280	2	Q5H9X5_ORYSA	Q5h9x5	oryza sativ	231	39	48.1	161	2	Q81SM3_BACAN	Q81sm3	bacillus an
159	40	49.4	286	2	Q93VV0_ARATH	Q93vv0	arabidopsis	232	39	48.1	179	2	Q8Y1M2_RALSO	Q8ylm2	ralstonia s
160	40	49.4	287	2	Q9SR31_ARATH	Q9sr31	arabidopsis	233	39	48.1	191	2	Q6MKZ6_BDEBA	Q6mkz6	bdellovibri
161	40	49.4	304	2	Q9RKV0_STRCO	Q9rkv0	streptomyce	234	39	48.1	195	2	O15640_LEICH	O15640	leishmania
162	40	49.4	313	2	Q8W0X5_MAIZE	Q8w0x5	zea mays (m	235	39	48.1	207	2	Q5DQH2_ECOLI	Q5dqh2	escherichia
163	40	49.4	351	2	Q76YE6_9CAUD	Q76ye6	bacterioph	236	39	48.1	208	2	Q68665_9GAMM	Q68665	cycloclasti
164	40	49.4	363	2	Q5BK01_RAT	Q5bk01	rattus norv	237	39	48.1	266	2	P74981_YEREN	P74981	yersinia en
165	40	49.4	372	2	O81073_ARATH	O81073	arabidopsis	238	39	48.1	298	2	Q6NFA4_CORDI	Q6nf44	corynebacte
166	40	49.4	373	2	Q67985_RHOOP	Q67985	rhodococcus	239	39	48.1	306	2	O81859_ARATH	O81859	arabidopsis
167	40	49.4	374	2	Q850U6_ORYSA	Q850u6	oryza sativ	240	39	48.1	308	2	Q81Q46_BACAN	Q81q46	bacillus an
168	40	49.4	378	2	Q9MNX0_PAGLO	Q9mnx0	pagurus lon	241	39	48.1	308	2	Q63AZ7_BACCCZ	Q63az7	bacillus ce
169	40	49.4	379	2	Q55JK6_CRYNE	Q55jk6	cryptococcu	242	39	48.1	324	1	O10R2_HUMAN	Q8ngx6	homo sapien
170	40	49.4	379	2	Q5K9Z5_CRYNE	Q5k9z5	cryptococcu	243	39	48.1	331	2	Q4JU09_CORJK	Q4ju09	corynebacte
171	40	49.4	393	2	Q97A07_THEVO	Q97a07	thermoplasm	244	39	48.1	334	2	Q9KWL7_CORGL	Q9kw17	corynebacte
172	40	49.4	398	2	Q5YBA1_9SPHN	Q5yba1	sphingomona	245	39	48.1	334	2	Q7BN99_PSEPU	Q7bn99	pseudomonas
173	40	49.4	410	1	ODBA_PSEPU	P09060	pseudomonas	246	39	48.1	334	2	Q88LX3_PSEPK	Q88lx3	pseudomonas
174	40	49.4	410	2	Q88EQ2_PSEPK	Q88eq2	pseudomonas	247	39	48.1	335	2	Q6IFS1_HUMAN	Q6ifs1	homo sapien
175	40	49.4	410	2	Q911M2_PSEAE	Q9ilm2	pseudomonas	248	39	48.1	356	2	Q6QMG3_ICTPU	Q6qmg3	ictalurus p
176	40	49.4	419	2	Q6ZTM6_HUMAN	Q6ztm6	homo sapien	249	39	48.1	374	2	Q6AJF4_DESPS	Q6ajf4	desulfotale
177	40	49.4	436	2	Q86VR7_HUMAN	Q86vr7	homo sapien	250	39	48.1	398	2	Q5SSY5_9HIV1	Q5s5y5	human immun

251	39	48.1	423	2	Q8AHK4_9HIV1	Q8ahk4	human immun	324	38	46.9	267	1	PENK_HUMAN	P01210	homo sapien
252	39	48.1	429	2	Q7UTX3_RHOBA	Q7utx3	rhodopirell	325	38	46.9	267	2	Q6FHC6_HUMAN	Q6fhc6	homo sapien
253	39	48.1	434	2	Q56AF2_9HIV1	Q56af2	human immun	326	38	46.9	267	2	Q6FHE6_HUMAN	Q6fhe6	homo sapien
254	39	48.1	456	2	Q526J5_MAGGR	Q526j5	magnaporthe	327	38	46.9	268	1	PENK_CAVPO	P47969	cavia porce
255	39	48.1	460	2	Q7WUA0_9GAMM	Q7wua0	cycloclasti	328	38	46.9	268	2	Q5ENN1_HETTR	Q5enn1	heterocapsa
256	39	48.1	465	2	Q9LZD8_ARATH	Q9lzd8	arabidopsis	329	38	46.9	275	2	Q66KB3_XENTR	Q66kb3	xenopus tro
257	39	48.1	481	2	Q51UK6_MAGGR	Q51uk6	magnaporthe	330	38	46.9	284	2	Q93UQ6_9SPHN	Q93uq6	agrobacteri
258	39	48.1	487	2	Q9N1R6_RABIT	Q9n1r6	oryctolagus	331	38	46.9	284	2	Q99H31_9NUCL	Q99h31	helicoverpa
259	39	48.1	498	2	Q7U3B1_SYNXP	Q7u3b1	synecchococc	332	38	46.9	286	2	Q92XB9_RHIME	Q92xb9	rhizobium m
260	39	48.1	501	2	Q67D59_9HIV1	Q67d59	human immun	333	38	46.9	290	2	Q7VZE1_BORPE	Q7vze1	bordetella
261	39	48.1	506	2	Q741T8_MYCPA	Q741t8	mycobacteri	334	38	46.9	290	2	Q7WGW7_BORBR	Q7wgw7	bordetella
262	39	48.1	510	2	Q4T2X4_TETNG	Q4t2x4	tetraodon n	335	38	46.9	293	2	Q6N2D7_RHOPA	Q6n2d7	rhodopseudo
263	39	48.1	517	2	Q9VB75_DROME	Q9vb75	prosoiphila	336	38	46.9	296	2	Q5FRV2_GLUOX	Q5frv2	gluconobact
264	39	48.1	529	2	Q4K961_PSEF5	Q4k961	pseudomonas	337	38	46.9	296	2	Q8DHQ3_SYNEL	Q8dhq3	synecchococc
265	39	48.1	579	2	Q5KDC7_CRYNE	Q5kdc7	cryptococcu	338	38	46.9	314	2	Q4ID97_GIBZE	Q4id97	gibberella
266	39	48.1	585	2	Q8VXA9_ORYSA	Q8vxa9	oryza sativ	339	38	46.9	315	2	Q87M69_VIBPA	Q87m69	vibrio para
267	39	48.1	595	1	LNT2_LEPIC	Q72md8	leptospira	340	38	46.9	318	2	Q6CR48_KLULA	Q6cr48	kluuveromyc
268	39	48.1	595	1	LNT2_LEPIN	Q8eyy4	leptospira	341	38	46.9	318	2	Q5N4R1_SYNP6	Q5n4r1	synecchococc
269	39	48.1	596	2	Q55Q40_CRYNE	Q55q40	cryptococcu	342	38	46.9	320	1	O2T33_HUMAN	Q8ng76	homo sapien
270	39	48.1	610	2	Q4IKG6_GIBZE	Q4ikg6	gibberella	343	38	46.9	320	1	OR2TC_HUMAN	Q8ng77	homo sapien
271	39	48.1	630	2	Q55ZM7_CRYNE	Q55zm7	cryptococcu	344	38	46.9	323	2	Q8FUJ8_ECOL6	Q8fjj8	escherichia
272	39	48.1	630	2	Q5KNZ0_CRYNE	Q5knz0	cryptococcu	345	38	46.9	323	2	Q6F982_ACTAD	Q6f982	acinetobact
273	39	48.1	637	2	Q4PGT8_USTMA	Q4pgt8	ustilago ma	346	38	46.9	329	2	Q8Q106_9HIV1	Q8q106	human immun
274	39	48.1	653	2	Q4QIU9_LEIMA	Q4qiu9	leishmania	347	38	46.9	329	2	Q8Q145_9HIV1	Q8q145	human immun
275	39	48.1	660	2	Q4S0V6_TETNG	Q4s0v6	tetraodon n	348	38	46.9	333	2	Q8LB62_ARATH	Q8lb62	arabidopsis
276	39	48.1	676	2	Q4QJ48_LEIMA	Q4qj48	leishmania	349	38	46.9	337	2	Q8RWW7_ARATH	Q8rww7	arabidopsis
277	39	48.1	735	1	DUR3_YEAST	Q33413	saccharomyc	350	38	46.9	342	2	Q9SVB8_ARATH	Q9svb8	arabidopsis
278	39	48.1	737	2	Q98IS7_RHILO	Q98is7	rhizobium l	351	38	46.9	352	2	Q801G3_BRARE	Q801g3	brachydanio
279	39	48.1	859	2	Q69RI8_ORYSA	Q69ri8	oryza sativ	352	38	46.9	352	2	Q568W4_BRARE	Q568w4	brachydanio
280	39	48.1	944	2	Q82QU4_STRAW	Q82qu4	streptomyce	353	38	46.9	357	2	O63702_9NEOP	O63702	panorpa jap
281	39	48.1	987	2	Q7YZM8_CAEEL	Q7yzm8	caenorhabdi	354	38	46.9	359	2	Q91BY5_9NUCL	Q91by5	helicoverpa
282	39	48.1	1003	2	Q7RSV9_PLAYO	Q7rsv9	plasmodium	355	38	46.9	361	2	Q8ZUX8_PYRAE	Q8zux8	pyrobaculum
283	39	48.1	1086	2	Q938K3_9CAUD	Q938k3	temperate p	356	38	46.9	362	2	O35281_MOUSE	O35281	mus musculu
284	39	48.1	1086	2	Q9A0N0_STRPY	Q9a0n0	streptococc	357	38	46.9	365	2	Q96L07_HUMAN	Q96l07	homo sapien
285	39	48.1	1086	2	Q8NZS4_STRP8	Q8nzs4	streptococc	358	38	46.9	365	2	Q85Q59_9HYST	Q85q59	thrichomys
286	39	48.1	1086	2	Q8K6Q1_STRP3	Q8k6q1	streptococc	359	38	46.9	365	2	Q85Q59_9HYST	Q85q59	thrichomys
287	39	48.1	1093	2	Q5XAS7_STRP6	Q5xas7	streptococc	360	38	46.9	370	1	CBNB_ALCEU	P83763	alcaligenes
288	39	48.1	1710	2	Q4RM05_TETNG	Q4rm05	tetraodon n	361	38	46.9	370	1	TFDD_COMAC	Q9rnz9	comamonas a
289	39	48.1	1730	2	Q7NK50_GLOVI	Q7nk50	gloeobacter	362	38	46.9	370	2	Q8GFH9_ALCXX	Q8gfh9	alcaligenes
290	39	48.1	1799	2	Q5KQE8_ORYSA	Q5kqe8	oryza sativ	363	38	46.9	370	2	Q9L3W6_PSECL	Q9l3w6	pseudomonas
291	39	48.1	1921	2	Q5JNB3_ORYSA	Q5jnb3	oryza sativ	364	38	46.9	374	2	Q7V818_PROMM	Q7v818	prochloroco
292	39	48.1	1939	2	Q5CRH8_CRYPV	Q5crh8	cryptospori	365	38	46.9	374	2	Q85Q47_9HYST	Q85q47	thrichomys
293	39	48.1	1969	2	Q852C7_ORYSA	Q852c7	oryza sativ	366	38	46.9	378	2	Q70EH4_9RODE	Q70eh4	saccostomus
294	39	48.1	2071	2	Q8GZY7_ORYSA	Q8gzy7	oryza sativ	367	38	46.9	378	2	Q70EH3_9RODE	Q70eh3	saccostomus
295	39	48.1	2145	2	Q8H7T1_ORYSA	Q8h7t1	oryza sativ	368	38	46.9	379	1	CYB_PROCY	Q9g2u7	proechimys
296	39	48.1	2256	2	Q7XPI7_ORYSA	Q7xpi7	oryza sativ	369	38	46.9	379	2	Q65Z02_9MAXI	Q65z02	megabalanus
297	39	48.1	2526	2	Q4WUP7_ASPFU	Q4wup7	aspergillus	370	38	46.9	379	2	Q85Q58_9HYST	Q85q58	thrichomys
298	38.5	47.5	179	2	Q6DIM7_XENTR	Q6dim7	xenopus tro	371	38	46.9	379	2	Q6XVQ6_PROCY	Q6xvq6	proechimys
299	38.5	47.5	509	2	Q50W76_ENTHI	Q50w76	entamoeba h	372	38	46.9	379	2	Q6XVQ5_PROCY	Q6xvq5	proechimys
300	38	46.9	106	2	Q4TTE9_CAERE	Q4tte9	caenorhabdi	373	38	46.9	379	2	Q9G0R5_PROCY	Q9g0r5	proechimys
301	38	46.9	111	2	Q54XP7_DICDI	Q54xp7	dictyosteli	374	38	46.9	380	2	Q508J4_9RODE	Q508j4	perognathus
302	38	46.9	112	2	P73045_SYNY3	P73045	synecchocyst	375	38	46.9	380	2	Q52CS2_MAGGR	Q52cs2	magnaporthe
303	38	46.9	126	2	Q74JP5_LACJO	Q74jp5	lactobacill	376	38	46.9	381	2	Q5TMD2_9CHOR	Q5tmd2	epigonichth
304	38	46.9	128	2	Q8PN10_XANAC	Q8pn10	xanthomonas	377	38	46.9	381	2	Q954G3_9RODE	Q954g3	sigmodon to
305	38	46.9	133	2	Q7TLB9_9VIRU	Q7tlb9	marine birn	378	38	46.9	381	2	Q8HHP2_9RODE	Q8hhp2	sigmodon to
306	38	46.9	154	2	Q62UM8_HUMAN	Q6zum8	homo sapien	379	38	46.9	381	2	Q8HHP1_9RODE	Q8hhp1	sigmodon to
307	38	46.9	166	2	Q96JS4_HUMAN	Q96js4	homo sapien	380	38	46.9	381	2	Q8HHP0_9RODE	Q8hhp0	sigmodon to
308	38	46.9	172	2	Q4TNN5_9SPHN	Q4tnn5	erythrobact	381	38	46.9	381	2	Q8HC70_9RODE	Q8hc70	sigmodon to
309	38	46.9	181	2	Q5K261_GUITH	Q5k261	guillardia	382	38	46.9	392	2	Q5FR07_GLUOX	Q5fr07	gluconobact
310	38	46.9	219	1	PIGF_HUMAN	Q07326	homo sapien	383	38	46.9	408	2	Q8DGS3_SYNEL	Q8dgs3	synecchococc
311	38	46.9	219	1	PIGF_MOUSE	O09101	mus musculu	384	38	46.9	410	2	Q5F493_CHICK	Q5f493	gallus gall
312	38	46.9	219	2	Q6IB04_HUMAN	Q6ib04	homo sapien	385	38	46.9	411	2	Q6P8K5_MOUSE	Q6p8k5	mus musculu
313	38	46.9	219	2	Q8K338_MOUSE	Q8k338	mus musculu	386	38	46.9	418	2	Q5W910_LINUN	Q5w910	lingula ung
314	38	46.9	221	2	Q8M2E3_9SAUR	Q8m2e3	pachydactyl	387	38	46.9	419	2	Q7NYR8_CHRVO	Q7nyr8	chromobacte
315	38	46.9	238	2	Q6XNZ4_9SAUR	Q6xnz4	phelsuma oc	388	38	46.9	435	2	Q51S05_MAGGR	Q51s05	magnaporthe
316	38	46.9	240	2	Q7U9H9_SYNXP	Q7u9h9	synecchococc	389	38	46.9	436	1	CYB_RHOCA	P08502	rhodobacter
317	38	46.9	243	2	Q8MLV9_DROME	Q8mlv9	drosoiphila	390	38	46.9	438	2	Q6A8A8_PROAC	Q6a8a8	propionibac
318	38	46.9	247	1	YCIC_ECO57	Q8xcb7	escherichia	391	38	46.9	441	2	O27107_METTH	O27107	methanobact
319	38	46.9	247	1	YCIC_ECOL6	Q8fhw3	escherichia	392	38	46.9	441	2	Q8DBK1_VIBVU	Q8dbk1	vibrio vuln
320	38	46.9	247	1	YCIC_ECOLI	P21365	escherichia	393	38	46.9	446	2	Q98ER5_RHILO	Q98er5	rhizobium l
321	38	46.9	247	1	YCIC_SHIFL	Q831c9	shigella fl	394	38	46.9	454	2	Q7PTI1_ANOGA	Q7pti1	anopheles g
322	38	46.9	247	2	Q9MP34_9HYME	Q9mp34	bothriomyrm	395	38	46.9	455	2	Q8ZDN3_YERPE	Q8zdn3	yersinia pe
323	38	46.9	247	2	Q8HZR8_MACMU	Q8hizr8	macaca mula	396	38	46.9	457	2	Q8D0S5_YERPE	Q8d0s5	yersinia pe

397	38	46.9	457	2	Q669C6_YERPS	Q669c6 yersinia ps	470	37	45.7	158	1	SMG_COXBU	Q83aa6 coxiella bu
398	38	46.9	464	2	Q77NK0_9GAMA	Q77nk0 rhesus monk	471	37	45.7	166	2	Q8SG46_9DIPT	Q8sg46 gasterophil
399	38	46.9	464	2	Q9WRS0_9GAMA	Q9wrs0 macaca mula	472	37	45.7	175	2	Q9B4A2_ANOST	Q9b4a2 anopheles s
400	38	46.9	465	2	Q6DHC7_BRARE	Q6dhc7 brachydanio	473	37	45.7	176	2	Q5EB63_RAT	Q5eb63 rattus norv
401	38	46.9	466	2	Q7MIC1_VIBVY	Q7mic1 vibrio vuln	474	37	45.7	182	2	Q702M8_9DIPT	Q702m8 anopheles f
402	38	46.9	472	2	Q8YH77_BRUME	Q8yh77 brucella me	475	37	45.7	182	2	Q705A4_9DIPT	Q705a4 anopheles j
403	38	46.9	474	2	Q4TDT8_TETNG	Q4tdt8 tetraodon n	476	37	45.7	190	2	Q61VC8_CAEBR	Q61vc8 caenorhabdi
404	38	46.9	477	2	Q8PI71_XANAC	Q8pi71 xanthomonas	477	37	45.7	192	2	Q4T8L5_TETNG	Q4t8l5 tetraodon n
405	38	46.9	486	2	Q5N9S1_ORYSA	Q5n9s1 oryza sativ	478	37	45.7	195	2	Q71S88_LEIDO	Q71s88 leishmania
406	38	46.9	489	2	Q8FC58_ECOL6	Q8fc58 escherichia	479	37	45.7	195	2	Q71S89_LEICH	Q71s89 leishmania
407	38	46.9	493	2	Q9LYK2_ARATH	Q9lyk2 arabidopsis	480	37	45.7	203	2	Q6TLC5_APIME	Q6tlc5 apis mellif
408	38	46.9	493	2	Q4LUX4_9BURK	Q4lux4 burkholderi	481	37	45.7	208	2	Q71S91_LEIDO	Q71s91 leishmania
409	38	46.9	499	2	Q9HZ88_PSEAE	Q9hz88 pseudomonas	482	37	45.7	208	2	Q71S92_LEICH	Q71s92 leishmania
410	38	46.9	504	2	Q4SDT4_TETNG	Q4sdt4 tetraodon n	483	37	45.7	220	2	Q85NU0_9SAUR	Q85nu0 pachydactyl
411	38	46.9	507	2	Q4KF18_PSEF5	Q4kfl8 pseudomonas	484	37	45.7	221	2	Q8M2D8_9SAUR	Q8m2d8 pachydactyl
412	38	46.9	511	2	Q983P8_RHILO	Q983p8 rhizobium l	485	37	45.7	225	2	Q4IYZ6_AZOVI	Q4iyz6 azotobacter
413	38	46.9	513	2	Q8GON0_BRUSU	Q8gon0 brucella su	486	37	45.7	225	2	Q88EG2_PSEPK	Q88eg2 pseudomonas
414	38	46.9	516	1	LNT_CHRVO	Q7nql3 chromobacte	487	37	45.7	239	2	Q7QWJ0_GIALA	Q7qwj0 giardia lam
415	38	46.9	516	2	Q94JG1_ORYSA	Q94jg1 oryza sativ	487	37	45.7	239	2	Q9LH11_ARATH	Q9lh11 arabidopsis
416	38	46.9	516	2	Q88LF8_PSEPK	Q88lf8 pseudomonas	489	37	45.7	248	2	Q9HWC1_PSEAE	Q9hwc1 pseudomonas
417	38	46.9	517	2	Q74D39_GEOSL	Q74d39 geobacter s	490	37	45.7	250	2	Q53U42_LYCES	Q53u42 lycopersico
418	38	46.9	526	2	Q7S3U3_NEUCR	Q7s3u3 neurospora	491	37	45.7	250	2	Q8WLA5_PETHY	Q8wla5 petunia hyb
419	38	46.9	529	2	Q9Z0E1_MOUSE	Q9z0e1 mus musculu	492	37	45.7	250	2	Q9SM37_SPOST	Q9sm37 sporobolus
420	38	46.9	530	2	Q8TC57_HUMAN	Q8tc57 homo sapien	493	37	45.7	253	2	Q7PMK2_ANOGA	Q7pmk2 anopheles g
421	38	46.9	533	2	Q7F952_ORYSA	Q7f952 oryza sativ	494	37	45.7	257	2	Q5PX07_HCMV	Q5px07 human cytom
422	38	46.9	533	2	Q5E6H2_VIBF1	Q5e6h2 vibrio fisc	495	37	45.7	260	2	Q49657_MYCLE	Q49657 mycobacteri
423	38	46.9	539	2	Q8EJE2_SHEON	Q8eje2 shewanella	496	37	45.7	266	1	HMUV_YERPE	Q56993 yersinia pe
424	38	46.9	541	2	Q32182_BACSU	Q32182 bacillus su	497	37	45.7	266	2	Q623V0_CAEBR	Q623v0 caenorhabdi
425	38	46.9	549	2	Q6U8A5_HUMAN	Q6u8a5 homo sapien	498	37	45.7	266	2	Q66FK0_YERPS	Q66fk0 yersinia ps
426	38	46.9	550	2	Q4RQ38_TETNG	Q4rq38 tetraodon n	499	37	45.7	272	2	Q8LAQ8_ARATH	Q8laq8 arabidopsis
427	38	46.9	580	2	Q8IVG3_HUMAN	Q8ivg3 homo sapien	500	37	45.7	274	2	Q984A1_RHILO	Q984a1 rhizobium l
428	38	46.9	645	2	Q97448_GIALA	Q97448 giardia lam	501	37	45.7	284	1	XTH13_ARATH	Q9fk18 arabidopsis
429	38	46.9	649	2	Q61M52_CAEBR	Q61m52 caenorhabdi	502	37	45.7	291	2	Q58PQ7_9PROT	Q58pq7 uncultured
430	38	46.9	668	2	Q965R5_CAEEL	Q965r5 caenorhabdi	503	37	45.7	295	2	Q6UKY2_9BURK	Q6uky2 burkholderi
431	38	46.9	673	2	Q45509_CAEEL	Q45509 caenorhabdi	504	37	45.7	297	1	YCEJ_ECOLI	P75845 escherichia
432	38	46.9	684	2	Q61AW5_CAEBR	Q61aw5 caenorhabdi	505	37	45.7	297	2	Q57R08_SALCH	Q57r08 salmonella
433	38	46.9	718	2	Q9RWG1_DEIRA	Q9rwg1 deinococcus	506	37	45.7	297	2	Q5PGF6_SALPA	Q5pgf6 salmonella
434	38	46.9	737	2	Q522D6_MAGGR	Q522d6 magnaporthe	507	37	45.7	297	2	Q8FJA8_ECOL6	Q8fja8 escherichia
435	38	46.9	790	2	Q96MR9_HUMAN	Q96mr9 homo sapien	508	37	45.7	297	2	Q8Z7Z9_SALTI	Q8z7z9 salmonella
436	38	46.9	796	2	Q5AMI8_CANAL	Q5ami8 candida alb	509	37	45.7	297	2	Q8ZQB9_SALTY	Q8zqb9 salmonella
437	38	46.9	799	1	SYFB_PARUW	Q5mea6 parachlamyd	510	37	45.7	299	2	Q6UKY1_9BURK	Q6uky1 burkholderi
438	38	46.9	854	1	DRL5_ARATH	Q9c8k0 arabidopsis	511	37	45.7	299	2	Q9PVL6_CHICK	Q9pvl6 gallus gall
439	38	46.9	855	1	POT13_ARATH	Q8lpl8 arabidopsis	512	37	45.7	301	2	Q63JA5_BURPS	Q63ja5 burkholderi
440	38	46.9	869	2	Q5L6Y5_CHLAB	Q5l6y5 chlamydophi	513	37	45.7	307	2	Q6UKX9_9BURK	Q6ukx9 burkholderi
441	38	46.9	1077	2	Q7ZZA5_BRARE	Q7zza5 brachydanio	514	37	45.7	307	2	Q6UKY0_9BURK	Q6uky0 burkholderi
442	38	46.9	1110	2	Q8HY22_PIG	Q8hy22 sus scrofa	515	37	45.7	310	1	OR4KE_HUMAN	Q8ngd5 homo sapien
443	38	46.9	1116	2	Q07898_HUMAN	Q07898 homo sapien	516	37	45.7	316	2	Q4K7V4_PSEF5	Q4k7v4 pseudomonas
444	38	46.9	1149	2	Q07901_HUMAN	Q07901 homo sapien	517	37	45.7	319	1	MGRD_RAT	Q7tn41 rattus norv
445	38	46.9	1151	2	Q07899_HUMAN	Q07899 homo sapien	518	37	45.7	330	2	Q53RL9_ORYSA	Q53rl9 oryza sativ
446	38	46.9	1156	2	Q86VB7_HUMAN	Q86vb7 homo sapien	519	37	45.7	333	2	Q926G3_RHIME	Q926g3 rhizobium m
447	38	46.9	1156	2	Q07900_HUMAN	Q07900 homo sapien	520	37	45.7	338	2	Q825J4_STRAW	Q825j4 streptomyce
448	38	46.9	1337	2	Q4PA65_USTMA	Q4pa65 ustilago ma	521	37	45.7	341	2	Q85700_STRLI	Q85700 streptomyce
449	38	46.9	1410	2	Q51L78_MAGGR	Q51l78 magnaporthe	522	37	45.7	341	2	Q7AKR4_STRCO	Q7akr4 streptomyce
450	38	46.9	1545	1	MRP2_HUMAN	Q92887 homo sapien	523	37	45.7	344	2	Q5AHZ3_CANAL	Q5ahz3 candida alb
451	38	46.9	1545	2	Q5T2B1_HUMAN	Q5t2b1 homo sapien	524	37	45.7	371	2	Q9LDP1_ARATH	Q9ldp1 arabidopsis
452	38	46.9	3165	2	Q5AR63_EMENI	Q5ar63 aspergillus	525	37	45.7	377	2	Q82PX5_STRAW	Q82px5 streptomyce
453	38	46.9	3618	2	Q4JTA1_CORJK	Q4jta1 corynebacte	526	37	45.7	378	2	Q91I23_PSEAE	Q91lz3 pseudomonas
454	37.5	46.3	336	2	Q6APK5_DESPS	Q6apk5 desulfotale	527	37	45.7	378	2	Q7U759_SYNPX	Q7u759 synchococc
455	37.5	46.3	509	2	Q8GM15_STRGL	Q8gm15 streptomyce	528	37	45.7	380	2	Q66SY3_9SAUR	Q66sy3 rhineura fl
456	37.5	46.3	571	1	PRIM_DEIRA	Q9rwr5 deinococcus	529	37	45.7	387	2	Q9Y7Y4_SCHPO	Q9y7y4 schizosacch
457	37.5	46.3	627	2	Q4T662_TETNG	Q4t662 tetraodon n	530	37	45.7	389	2	Q5FHU7_LACAC	Q5fhu7 lactobacill
458	37	45.7	80	2	Q84Z31_ORYSA	Q84z31 oryza sativ	531	37	45.7	390	2	Q90I59_9HIV1	Q90i59 human immun
459	37	45.7	82	2	Q71IQ2_LACDL	Q71iq2 lactobacill	532	37	45.7	391	2	Q4J0D5_AZOVI	Q4j0d5 azotobacter
460	37	45.7	86	2	Q98NI4_RHILO	Q98ni4 rhizobium l	533	37	45.7	395	2	Q5SRJ5_HUMAN	Q5srj5 homo sapien
461	37	45.7	92	2	Q76YD3_9CAUD	Q76yd3 bacterioph	534	37	45.7	401	1	MDTH_CANBF	Q7vqx5 candidatus
462	37	45.7	99	2	Q8ANJ3_9HIV1	Q8anj3 human immun	535	37	45.7	401	2	Q90IU9_9HIV1	Q90iu9 human immun
463	37	45.7	121	2	Q87KN7_VIBPA	Q87kn7 vibrio para	536	37	45.7	403	2	Q88NU2_PSEPK	Q88nu2 pseudomonas
464	37	45.7	123	2	Q4TRR4_9SPHN	Q4trr4 erythroba	537	37	45.7	410	2	Q4LN26_9BURK	Q4ln26 burkholderi
465	37	45.7	128	2	Q9GAB2_MERCA	Q9gab2 mercenaria	538	37	45.7	410	2	Q5WTU9_LEGPL	Q5wtu9 legionella
466	37	45.7	134	2	Q9MC43_BPD3	Q9mc43 bacterioph	539	37	45.7	410	2	Q5X223_LEGPA	Q5x223 legionella
467	37	45.7	147	2	Q6ZRM0_HUMAN	Q6zrm0 homo sapien	540	37	45.7	410	2	Q5ZSI6_LEGPH	Q5zsl6 legionella
468	37	45.7	150	2	Q8TS13_METAC	Q8tsl3 methanosarc	541	37	45.7	410	2	Q6JHZ6_BURPS	Q6jhz6 burkholderi
469	37	45.7	150	2	Q5MZI4_SYNPE	Q5mzi4 synechococc	542	37	45.7	410	2	Q629V1_BURMA	Q629v1 burkholderi

689	36	44.4	77	2	Q5PAK1_ANAMM	Q5pak1 anaplasma m	762	36	44.4	267	2	Q8KG89_CHLTE	Q8kg89 chlorobium
690	36	44.4	77	2	Q5P9H0_ANAMM	Q5p9h0 anaplasma m	763	36	44.4	269	2	Q6NHR3_CORDI	Q6nhr3 corynebacte
691	36	44.4	77	2	Q5PBV7_ANAMM	Q5pbv7 anaplasma m	764	36	44.4	269	2	Q67MB9_SYMTH	Q67mb9 symbiobacte
692	36	44.4	78	2	Q7WLR9_BORBR	Q7wlr9 bordetella	765	36	44.4	273	2	Q9D8H4_MOUSE	Q9d8h4 mus musculu
693	36	44.4	83	2	Q82W04_NITEU	Q82w04 nitrosomona	766	36	44.4	277	2	Q53K02_ORYSA	Q53k02 oryza sativ
694	36	44.4	89	2	Q83TD1_ANAMA	Q83td1 anaplasma m	767	36	44.4	277	2	Q811P9_MOUSE	Q811f9 mus musculu
695	36	44.4	89	2	Q84H53_ANAMA	Q84h53 anaplasma m	768	36	44.4	290	2	Q82BH7_STRAW	Q82bh7 streptomyce
696	36	44.4	89	2	Q5P9V5_ANAMM	Q5p9v5 anaplasma m	769	36	44.4	295	2	O03320_9SAUR	O03320 elseya lati
697	36	44.4	91	2	Q84HV1_ANAMA	Q84hv1 anaplasma m	770	36	44.4	297	2	O03335_9SAUR	O03335 pelusios wi
698	36	44.4	99	2	Q9J079_9HIV1	Q9j079 human immun	771	36	44.4	298	2	Q6P519_HUMAN	Q6p519 homo sapien
699	36	44.4	102	2	Q4G0Q7_HUMAN	Q4g0q7 homo sapien	772	36	44.4	300	2	Q51RR3_MAGGR	Q51rr3 magnaporthe
700	36	44.4	110	2	Q05841_YEAST	Q05841 saccharomyc	773	36	44.4	300	2	Q628N2_CAEBR	Q628n2 caenorhabdi
701	36	44.4	118	2	Q60912_MOUSE	Q60912 mus musculu	774	36	44.4	301	2	Q8WY00_HUMAN	Q8wy00 homo sapien
702	36	44.4	128	2	Q7XV70_ORYSA	Q7xv70 oryza sativ	775	36	44.4	301	2	Q6AGK6_LEIXX	Q6agk6 leifsonia x
703	36	44.4	137	1	CRCB1_BACCI	P61386 bacillus ce	776	36	44.4	301	2	Q8C790_MOUSE	Q8c790 mus musculu
704	36	44.4	145	2	Q89DP5_BRAJA	Q89dp5 bradyrhizob	777	36	44.4	302	2	Q55VB5_CRYNE	Q55vb5 cryptococcu
705	36	44.4	147	2	Q865E3_CANPA	Q865e3 canis famil	778	36	44.4	302	2	Q5KXK0_CRYNE	Q5kkx0 cryptococcu
706	36	44.4	149	2	Q896D4_CLOTE	Q896d4 clostridium	779	36	44.4	304	2	Q8Z1K7_SALTI	Q8z1k7 salmonella
707	36	44.4	166	1	HA22G_ARATH	Q91r09 arabidopsis	780	36	44.4	305	2	Q5ZD45_ORYSA	Q5zd45 oryza sativ
708	36	44.4	173	2	Q7NPI3_GLOVI	Q7npi3 gloeobacter	781	36	44.4	308	1	GSHB_PROMA	Q7tvb0 prochloroco
709	36	44.4	174	2	Q5YZH1_NOCFA	Q5yzh1 nocardia fa	782	36	44.4	309	2	Q8NI49_HUMAN	Q8ni49 homo sapien
710	36	44.4	175	2	Q9B496_9DIPT	Q9b496 anopheles b	783	36	44.4	309	2	Q9JKT8_RAT	Q9jkt8 rattus norv
711	36	44.4	175	2	Q9B497_9DIPT	Q9b497 anopheles c	784	36	44.4	313	1	COFG_SYNY3	P73191 synechocyst
712	36	44.4	175	2	Q9B4A5_ANOAL	Q9b4a5 anopheles a	785	36	44.4	315	2	Q34185_CEPNE	Q34185 cepaea nemo
713	36	44.4	175	2	Q9B4A6_ANOAB	Q9b4a6 anopheles a	786	36	44.4	318	2	Q53J43_ORYSA	Q53j43 oryza sativ
714	36	44.4	182	2	Q702M9_9DIPT	Q702m9 anopheles a	787	36	44.4	319	2	Q5UF78_9PROT	Q5uf78 uncultured
715	36	44.4	182	2	Q702N0_9DIPT	Q702n0 anopheles a	788	36	44.4	320	2	Q4UZV7_XANCP	Q4uzv7 xanthomonas
716	36	44.4	182	2	Q702N2_9DIPT	Q702n2 anopheles a	789	36	44.4	320	2	Q8PDL6_XANCP	Q8pdl6 xanthomonas
717	36	44.4	182	2	Q702N3_9DIPT	Q702n3 anopheles a	790	36	44.4	321	2	Q4KEA0_PSEF5	Q4kea0 pseudomonas
718	36	44.4	182	2	Q702N5_9DIPT	Q702n5 anopheles a	791	36	44.4	323	2	Q7U936_SYNPX	Q7u936 synechococc
719	36	44.4	192	2	Q6N4D2_RHOPA	Q6n4d2 rhodopseudo	792	36	44.4	324	1	TTP_BOVIN	P53781 bos taurus
720	36	44.4	195	2	O02616_TRYCR	O02616 trypanosoma	793	36	44.4	326	2	Q5BJY7_RAT	Q5bjy7 rattus norv
721	36	44.4	195	2	Q27791_TRYCR	Q27791 trypanosoma	794	36	44.4	330	1	A85B_MYCAV	Q06947 mycobacteri
722	36	44.4	195	2	Q71S87_LEIMA	Q71s87 leishmania	795	36	44.4	332	2	Q8KIQ1_PSEAE	Q8kiq1 pseudomonas
723	36	44.4	195	2	Q4Q597_LEIMA	Q4q597 leishmania	796	36	44.4	334	2	Q742M3_MYCPA	Q742m3 mycobacteri
724	36	44.4	199	2	Q87756_KLEPN	Q87756 klebsiella	797	36	44.4	335	2	Q9SZ96_ARATH	Q9sz96 arabidopsis
725	36	44.4	200	2	Q4HGX1_CAMCO	Q4hgx1 campylobact	798	36	44.4	339	2	Q4WN85_ASPFU	Q4wn85 aspergillus
726	36	44.4	201	2	Q5Y3J0_9NEOP	Q5y3j0 pedicinus h	799	36	44.4	344	2	Q7YHM3_9NEOP	Q7yhm3 lepidopsoci
727	36	44.4	204	2	P91208_CABEL	P91208 caenorhabdi	800	36	44.4	344	2	Q5PPG3_RAT	Q5ppg3 rattus norv
728	36	44.4	206	2	Q72D10_DESVH	Q72d10 desulfovibr	801	36	44.4	346	1	OXDA_RAT	O35078 rattus norv
729	36	44.4	210	2	Q8NCH4_HUMAN	Q8nch4 homo sapien	802	36	44.4	347	2	Q9PT80_NOTVI	Q9pt80 notophthalm
730	36	44.4	219	2	Q6D9S3_ERWCT	Q6d9s3 erwinia car	803	36	44.4	349	2	Q6CRV3_KLULA	Q6crv3 kluyveromyc
731	36	44.4	220	2	Q85NV1_9SAUR	Q85nv1 pachyductyl	804	36	44.4	359	2	Q8JTH5_9CLOS	Q8jth5 little cher
732	36	44.4	221	2	Q957I9_LACAG	Q957i9 lacerta agi	805	36	44.4	360	2	Q4NA72_9MICC	Q4na72 arthrobacte
733	36	44.4	221	2	Q8M2D2_9SAUR	Q8m2d2 pachyductyl	806	36	44.4	365	2	Q8WY01_HUMAN	Q8wy01 homo sapien
734	36	44.4	222	2	Q6SN77_9PRIM	Q6sn77 cercocebus	807	36	44.4	365	2	Q5YQK0_NOCFA	Q5yqk0 nocardia fa
735	36	44.4	223	2	Q7XJ22_ORYSA	Q7xj22 oryza sativ	808	36	44.4	370	1	TFDD1_ALCEU	P05404 alcaligenes
736	36	44.4	231	2	Q82MW7_STRAW	Q82mw7 streptomyce	809	36	44.4	371	2	Q8G9L1_RHOOP	Q8g9l1 rhodococcus
737	36	44.4	233	1	YC53L_SYNY3	P72583 synechocyst	810	36	44.4	373	2	Q8WXXZ9_HUMAN	Q8wxz9 homo sapien
738	36	44.4	233	2	Q6Z9K2_ORYSA	Q6z9k2 oryza sativ	811	36	44.4	379	2	Q585J6_9TRYP	Q585j6 trypanosoma
739	36	44.4	234	2	Q75CR4_ASHGO	Q75cr4 ashbya goss	812	36	44.4	381	2	Q954G4_9RODE	Q954g4 sigmodon le
740	36	44.4	235	2	Q6MPK0_BDEBA	Q6mpk0 bdellovibri	813	36	44.4	382	2	O99971_PORPU	O99971 porphyra pu
741	36	44.4	237	2	Q6XNT0_9SAUR	Q6xnt0 phelsuma gu	814	36	44.4	385	2	O6NG02_CORDI	O6ng02 corynebacte
742	36	44.4	238	2	Q9MP31_9HYME	Q9mp31 dorymyrmex	815	36	44.4	389	2	Q7QVC2_GIALA	Q7qvc2 giardia lam
743	36	44.4	238	2	Q6XNS9_9SAUR	Q6xns9 phelsuma gu	816	36	44.4	390	2	Q9R6K4_9RHIZ	Q9r6k4 agrobacteri
744	36	44.4	242	2	Q4T0E6_TETNG	Q4t0e6 tetraodon n	817	36	44.4	391	2	Q9TC99_NEPOL	Q9tc99 nephroselmi
745	36	44.4	243	2	Q6D1D9_ERWCT	Q6d1d9 erwinia car	818	36	44.4	393	2	Q5PBM2_ANAMM	Q5pbm2 anaplasma m
746	36	44.4	244	2	Q5YJH8_9SAUR	Q5yjh8 amphiglossu	819	36	44.4	394	2	Q88ZC8_LACPL	Q88zc8 lactobacill
747	36	44.4	249	2	Q4V8D7_RAT	Q4v8d7 rattus norv	820	36	44.4	398	2	Q7B5K6_LACPL	Q7b5k6 lactobacill
748	36	44.4	257	2	Q5PX01_HCMV	Q5px01 human cytom	821	36	44.4	400	2	Q4SGC7_TETNG	Q4sgc7 tetraodon n
749	36	44.4	257	2	Q5PX04_HCMV	Q5px04 human cytom	822	36	44.4	402	2	Q6BND0_DEBHA	Q6bnd0 debaryomyce
750	36	44.4	257	2	Q5PX09_HCMV	Q5px09 human cytom	823	36	44.4	406	2	Q5G534_PHOSU	Q5g534 phodopus su
751	36	44.4	257	2	Q52N52_HCMV	Q52n52 human cytom	824	36	44.4	409	2	Q93CU8_SHIBO	Q93cu8 shigella bo
752	36	44.4	257	2	Q52N53_HCMV	Q52n53 human cytom	825	36	44.4	410	2	Q8NTI5_CORGL	Q8nti5 corynebacte
753	36	44.4	257	2	Q52N54_HCMV	Q52n54 human cytom	826	36	44.4	413	2	Q541U0_RAT	Q541u0 rattus norv
754	36	44.4	257	2	Q52N55_HCMV	Q52n55 human cytom	827	36	44.4	417	2	Q52CF4_MAGGR	Q52cf4 magnaporthe
755	36	44.4	258	2	Q89Q37_BRAJA	Q89q37 bradyrhizob	828	36	44.4	423	2	Q528C4_MAGGR	Q528c4 magnaporthe
756	36	44.4	258	2	Q7TML5_MOUSE	Q7tml5 mus musculu	829	36	44.4	423	2	Q5G4F9_9HIV1	Q5g4f9 human immun
757	36	44.4	258	2	Q5PX08_HCMV	Q5px08 human cytom	830	36	44.4	423	2	Q5G4C6_9HIV1	Q5g4c6 human immun
758	36	44.4	258	2	Q5PX15_HCMV	Q5px15 human cytom	831	36	44.4	428	1	NEUR3_BOVIN	O97859 bos taurus
759	36	44.4	260	1	CCCP_DROME	O76879 drosophila	832	36	44.4	428	2	Q4SGC9_TETNG	Q4sgc9 tetraodon n
760	36	44.4	260	2	Q5YML6_NOCFA	Q5yml6 nocardia fa	833	36	44.4	434	2	Q53EU6_HUMAN	Q53eu6 homo sapien
761	36	44.4	264	2	Q5YWL2_NOCFA	Q5ywl2 nocardia fa	834	36	44.4	434	2	Q96NA3_HUMAN	Q96na3 homo sapien

835	36	44.4	439	2	Q9E1X0_9ALPH	Q9elx0 cercopithec	36	44.4	567	1	Y1354_MYCBO	P59972 mycobacteri
836	36	44.4	439	2	Q4TH55_TETNG	Q4th55 tetraodon n	36	44.4	567	2	Q4UDZ9_THEAN	Q4udz9 theileria a
837	36	44.4	442	2	Q52DK1_MAGGR	Q52dk1 magnaporth	36	44.4	568	2	Q8VCU1_MOUSE	Q8vcu1 mus musculu
838	36	44.4	444	2	Q60A44_METCA	Q60a44 methylococc	36	44.4	572	2	Q9XMN5_CERCA	Q9xmn5 ceratitis c
839	36	44.4	445	1	HRH3_CAVPO	Q9ji35 cavia porce	36	44.4	575	2	Q8YK00_ANASP	Q8yk00 anabaena sp
840	36	44.4	445	1	HRH3_HUMAN	Q9y5n1 homo sapien	36	44.4	575	2	Q8DXQ4_STRAS	Q8dxq4 streptococc
841	36	44.4	445	1	HRH3_MOUSE	P58406 mus musculu	36	44.4	575	2	Q8E3C3_STRAS	Q8e3c3 streptococc
842	36	44.4	445	1	HRH3_RAT	Q9qyn8 rattus norv	36	44.4	576	1	CATA_RHOCA	P37743 rhodobacter
843	36	44.4	445	2	Q548M6_HUMAN	Q9qyn8 rattus norv	36	44.4	580	2	Q5JCK2_AEDAL	Q5jck2 aedes albop
844	36	44.4	445	2	Q4QRI7_HUMAN	Q548m6 homo sapien	36	44.4	580	2	Q4ZST9_PSESY	Q4zst9 pseudomonas
845	36	44.4	445	2	Q865E1_MACMU	Q4qri7 homo sapien	36	44.4	584	2	Q7SEB6_NEUCR	Q7seb6 neurospora
846	36	44.4	445	2	Q540P3_MOUSE	Q865e1 macaca mula	36	44.4	586	2	Q6BTF0_DEBHA	Q6btf0 debaryomyce
847	36	44.4	445	2	Q5G535_PHOSU	Q540p3 mus musculu	36	44.4	588	2	Q41RB6_9BURK	Q41rb6 burkholderi
848	36	44.4	446	2	Q5F795_NEIG1	Q5g535 phodopus su	36	44.4	588	2	Q8YJV1_ANASP	Q8yjv1 anabaena sp
849	36	44.4	446	2	Q9JYD0_NEIMB	Q5f795 neisseria g	36	44.4	592	1	THD1_ARATH	Q9zss6 arabidopsis
850	36	44.4	447	2	Q5YYK2_NOCPA	Q9jyd0 neisseria m	36	44.4	592	2	Q8GUG5_ARATH	Q8gug5 arabidopsis
851	36	44.4	449	2	Q51XE0_MAGGR	Q5yyk2 nocardia fa	36	44.4	597	2	Q25512_MANSE	Q25512 manduca sex
852	36	44.4	449	2	Q8NN88_CORGL	Q51xe0 magnaporth	36	44.4	600	2	Q8CE00_MOUSE	Q8ce00 mus musculu
853	36	44.4	454	2	Q65AT1_9SPHN	Q8nn88 corynebacte	36	44.4	604	2	Q4UI33_THEAN	Q4ui33 theileria a
854	36	44.4	454	2	Q67RK9_SYMTH	Q65at1 sphingomona	36	44.4	606	2	Q6CC14_YARLI	Q6cc14 yarrowia li
855	36	44.4	455	1	GUDP_BACSU	Q67rk9 symbiobacte	36	44.4	615	2	Q93L36_RHILV	Q9njl0 trichoplusi
856	36	44.4	456	2	Q4KCE9_PSEF5	P42237 bacillus su	36	44.4	623	2	Q8BY18_MOUSE	Q4wq64 aspergillus
857	36	44.4	457	2	Q4JXV1_CORJK	Q4kce9 pseudomonas	36	44.4	627	2	Q8BY18_MOUSE	Q8by18 mus musculu
858	36	44.4	459	2	Q9C772_ARATH	Q9c772 arabidopsis	36	44.4	627	2	Q25272_LEIDO	Q25272 leishmania
859	36	44.4	460	2	Q51WV3_MAGGR	Q51wv3 magnaporth	36	44.4	629	2	Q7Q126_ANOGA	Q7q126 anopheles g
860	36	44.4	461	2	Q5KN29_CRYNE	Q5kn29 cryptococcu	36	44.4	631	2	Q25294_LEIIN	Q25294 leishmania
861	36	44.4	467	2	Q7SAC8_NEUCR	Q7sac8 neurospora	36	44.4	631	2	Q7Q126_ANOGA	Q7q126 anopheles g
862	36	44.4	473	2	Q6ZS72_HUMAN	Q6zsb72 homo sapien	36	44.4	634	2	Q7Q126_ANOGA	Q7q126 anopheles g
863	36	44.4	473	2	Q7UQU29_RHOBA	Q7uq29 rhodopirell	36	44.4	634	2	Q9N6S9_LEIMA	Q9n6s9 leishmania
864	36	44.4	474	2	Q972Y5_SULTO	Q972y5 sulfolobus	36	44.4	645	1	Q7Q4C7_ANOGA	Q7q4c7 anopheles g
865	36	44.4	474	2	Q51QN4_MAGGR	Q51qn4 magnaporth	36	44.4	645	1	OPGH_XANCP	Q8p532 xanthomonas
866	36	44.4	474	2	Q4I4F1_GIBZE	Q4i4f1 gibberella	36	44.4	645	2	Q4UYZ8_XANCP	Q4uyz8 xanthomonas
867	36	44.4	474	2	Q7QXE7_GIALA	Q7qxe7 giardia lam	36	44.4	645	2	Q5ZMQ9_CHICK	Q5zmq9 gallus gall
868	36	44.4	476	2	Q7QE72_ANOGA	Q7qe72 anopheles g	36	44.4	649	2	Q75G80_ORYSA	Q75g80 oryza sativ
869	36	44.4	477	2	Q9RUQ5_DEIRA	Q9rug5 deinococcu	36	44.4	650	2	Q55YQ8_CRYNE	Q55yg8 cryptococcu
870	36	44.4	478	2	Q6ZHP8_ORYSA	Q6zhp8 oryza sativ	36	44.4	668	1	YNG5_SCHPO	O13621 schizosacch
871	36	44.4	484	2	Q83R75_SHIFL	Q83r75 shigella fl	36	44.4	673	1	NAL10_MOUSE	Q8ccn1 mus musculu
872	36	44.4	488	2	Q50E62_MOUSE	Q50e62 mus musculu	36	44.4	673	2	Q6JGS9_MOUSE	Q6jgs9 mus musculu
873	36	44.4	491	2	Q8N115_HUMAN	Q8n115 homo sapien	36	44.4	676	1	UL06_HHV11	P10190 human herpe
874	36	44.4	494	2	Q7XH54_ORYSA	Q7xh54 oryza sativ	36	44.4	676	1	UL06_HHV2H	P89429 human herpe
875	36	44.4	494	2	Q8W5K4_ORYSA	Q8w5k4 oryza sativ	36	44.4	688	2	Q80ZE3_MOUSE	Q80ze3 mus musculu
876	36	44.4	498	2	Q9SZ45_ARATH	Q9sz45 arabidopsis	36	44.4	704	2	Q6PUM2_LEIIN	Q6pum2 leishmania
877	36	44.4	502	2	Q8BVF9_MOUSE	Q8bvf9 mus musculu	36	44.4	704	2	Q9UAC2_LEIDO	Q9uac2 leishmania
878	36	44.4	503	2	Q9V7R7_DROME	Q9v7r7 drosophila	36	44.4	708	2	Q4QHH9_LEIMA	Q4qhh9 leishmania
879	36	44.4	505	2	Q8NA69_HUMAN	Q8na69 homo sapien	36	44.4	709	2	Q53QB3_ORYSA	Q53qb3 oryza sativ
880	36	44.4	513	2	Q93SD6_ECOLI	Q93sd6 escherichia	36	44.4	711	2	Q51HJ9_MAGGR	Q51hj9 magnaporth
881	36	44.4	513	2	Q4FBH0_ECOLI	Q4fbh0 escherichia	36	44.4	713	2	Q4QHH7_LEIMA	Q4qhh7 leishmania
882	36	44.4	517	1	XYLB_BUTFI	P45982 butyrivibri	36	44.4	720	2	Q8MTI4_LEIMA	Q8mti4 leishmania
883	36	44.4	519	2	Q9N3G2_CAEEL	Q9n3g2 caenorhabdi	36	44.4	722	2	Q9UAB8_LEIDO	Q9uab8 leishmania
884	36	44.4	520	2	Q7QDI8_ANOGA	Q7qdi8 anopheles g	36	44.4	724	2	Q6FN32_CANGA	Q6fn32 candida gla
885	36	44.4	521	2	Q8H925_ORYSA	Q8h925 oryza sativ	36	44.4	725	2	Q7XH77_ORYSA	Q7xh77 oryza sativ
886	36	44.4	524	2	Q9KSI8_VIBCH	Q9ksi8 vibrio chol	36	44.4	725	2	Q8W5M3_ORYSA	Q8w5m3 oryza sativ
887	36	44.4	524	2	Q91XD5_MOUSE	Q91xd5 mus musculu	36	44.4	728	2	Q51J30_MAGGR	Q51j30 magnaporth
888	36	44.4	524	2	Q5ZLX4_CHICK	Q5zlx4 gallus gall	36	44.4	732	2	Q4QHH8_LEIMA	Q4qhh8 leishmania
889	36	44.4	527	2	Q51SD7_MAGGR	Q51sd7 magnaporth	36	44.4	749	2	Q93598_BRARE	Q93598 brachydanio
890	36	44.4	527	2	Q4M0E4_9BURK	Q4m0e4 burkholderi	36	44.4	749	2	Q6P943_BRARE	Q6p943 brachydanio
891	36	44.4	528	2	Q4UVT6_XANCP	Q4uvt6 xanthomonas	36	44.4	766	2	Q8C4J2_MOUSE	Q8c4j2 mus musculu
892	36	44.4	528	2	Q8P8A1_XANCP	Q8p8a1 xanthomonas	36	44.4	771	2	Q4V782_MOUSE	Q4v782 mus musculu
893	36	44.4	528	2	Q8PJQ6_XANAC	Q8pjq6 xanthomonas	36	44.4	775	2	Q7F9A6_ORYSA	Q7f9a6 oryza sativ
894	36	44.4	530	2	Q526K5_MAGGR	Q526k5 magnaporth	36	44.4	775	2	Q9SU76_ARATH	Q9su76 arabidopsis
895	36	44.4	530	2	Q5GZ38_XANOR	Q5gz38 xanthomonas	36	44.4	779	2	Q728Y0_DESVH	Q728y0 desulfovibr
896	36	44.4	535	2	Q51QM7_MAGGR	Q51qm7 magnaporth	36	44.4	786	2	Q75LD4_ORYSA	Q75ld4 oryza sativ
897	36	44.4	537	2	Q8XS18_RALSO	Q8xs18 ralstonia s	36	44.4	786	2	Q7XKS1_ORYSA	Q7xks1 oryza sativ
898	36	44.4	538	2	Q92U07_RHIME	Q92u07 rhizobium m	36	44.4	789	2	Q75DB5_ASHGO	Q75db5 ashbya goss
899	36	44.4	541	2	Q9K2X5_STRCO	Q9kzx5 streptomyce	36	44.4	816	2	Q4NTY9_9DELT	Q4nty9 anaeromyxob
900	36	44.4	545	2	Q4IF76_GIBZE	Q4ift6 gibberella	36	44.4	838	2	Q8L250_MOUSE	Q8l250 mus musculu
901	36	44.4	551	2	Q7VMW0_HAEDU	Q7vmw0 haemophilus	36	44.4	845	2	Q8L4Q1_ORYSA	Q8l4q1 oryza sativ
902	36	44.4	557	2	Q51QW7_MAGGR	Q51qw7 magnaporth	36	44.4	852	2	Q6PB55_MOUSE	Q6pb55 mus musculu
903	36	44.4	558	1	Y1145_VIBCH	Q9ksw1 vibrio chol	36	44.4	857	2	Q4NN47_9DELT	Q4nn47 anaeromyxob
904	36	44.4	558	2	Q9V6X6_DROME	Q9v6x6 drosophila	36	44.4	885	1	ATG9_PICPA	Q876n4 pichia past
905	36	44.4	560	1	Y2420_PHOPR	Q6lph2 photobacter	36	44.4	922	2	Q8NAZ1_HUMAN	Q8naz1 homo sapien
906	36	44.4	562	2	Q4Q3B8_LEIMA	Q4q3b8 leishmania	36	44.4	923	2	Q4P3M5_USTMA	Q4p3m5 ustilago ma
907	36	44.4	567	1	Y1320_MYCTU	Q10633 mycobacteri	36	44.4	927	2	Q4YZ18_PLABE	Q4yz18 plasmodium

Y1354_MYCBO	P59972 mycobacteri
Q4UDZ9_THEAN	Q4udz9 theileria a
Q8VCU1_MOUSE	Q8vcu1 mus musculu
Q9XMN5_CERCA	Q9xmn5 ceratitis c
Q8YK00_ANASP	Q8yk00 anabaena sp
Q8DXQ4_STRAS	Q8dxq4 streptococc
Q8E3C3_STRAS	Q8e3c3 streptococc
CATA_RHOCA	P37743 rhodobacter
Q5JCK2_AEDAL	P57432 aedes albop
Q4ZST9_PSESY	Q4zst9 pseudomonas
Q7SEB6_NEUCR	Q7seb6 neurospora
Q6BTF0_DEBHA	Q6btf0 debaryomyce
Q4LRB6_9BURK	Q4lrb6 burkholderi
Q8YJV1_ANASP	Q8yjv1 anabaena sp
THD1_ARATH	Q9zss6 arabidopsis
Q8GUG5_ARATH	Q8gug5 arabidopsis
Q25512_MANSE	Q25512 manduca sex
Q8CE00_MOUSE	Q8ce00 mus musculu
Q4UI33_THEAN	Q4ui33 theileria a
Q6CC14_YARLI	Q6cc14 yarrowia li
Q9NJ10_TRINI	Q9njl0 trichoplusi
Q4WQ64_ASPFU	Q4wq64 aspergillus
Q93L36_RHILV	Q93l36 rhizobium l
Q8BY18_MOUSE	Q8by18 mus musculu
Q25272_LEIDO	Q25272 leishmania
Q25294_LEIIN	Q25294 leishmania
Q7Q126_ANOGA	Q7q126 anopheles g
Q76729_LEIME	Q76729 leishmania
Q9N6S9_LEIMA	Q9n6s9 leishmania
Q7Q4C7_ANOGA	Q7q4c7 anopheles g
OPGH_XANCP	Q8p532 xanthomonas
Q4UYZ8_XANCP	Q4uyz8 xanthomonas
Q5ZMQ9_CHICK	Q5zmq9 gallus galli
Q75G80_ORYSA	Q75g80 oryza sativa
Q55YQ8_CRYNE	Q55yq8 cryptococcus
YNG5_SCHPO	O13621 schizosacchar
NAL10_MOUSE	Q8ccn1 mus musculu
Q6JGS9_MOUSE	Q6jgs9 mus musculu
UL06_HHV11	P10190 human herpes
UL06_HHV2H	P89429 human herpes
Q80ZE3_MOUSE	Q80ze3 mus musculu
Q6PUM2_LEIIN	Q6pum2 leishmania
Q9UAC2_LEIDO	Q9uac2 leishmania
Q4QHH9_LEIMA	Q4qhh9 leishmania
Q53QB3_ORYSA	Q53qb3 oryza sativa
Q51HJ9_MAGGR	Q51hj9 magnaporthe
Q4QHH7_LEIMA	Q4qhh7 leishmania
Q8MT14_LEITA	Q8mti4 leishmania
Q9UAB8_LEIDO	Q9uab8 leishmania
Q6FN32_CANGA	Q6fn32 candida glabr
Q7XH77_ORYSA	Q7xh77 oryza sativa
Q8W5M3_ORYSA	Q8w5m3 oryza sativa
Q51J30_MAGGR	Q51j30 magnaporthe
Q4QHH8_LEIMA	Q4qhh8 leishmania
Q93598_BRARE	Q93598 brachydanio
Q6P943_BRARE	Q6p943 brachydanio
Q8C4J2_MOUSE	Q8c4j2 mus musculu
Q4V782_MOUSE	Q4v782 mus musculu
Q7F9A6_ORYSA	Q7f9a6 oryza sativa
Q9SU76_ARATH	Q9su76 arabidopsis
Q728Y0_DESVH	Q728y0 desulfovibrio
Q75LD4_ORYSA	Q75ld4 oryza sativa
Q7XKS1_ORYSA	Q7xks1 oryza sativa
Q75DB5_ASHGO	Q75db5 ashbya goss
Q4NTY9_9DELT	Q4nty9 anaeromyxob
Q8K250_MOUSE	Q8k250 mus musculu
Q8L4Q1_ORYSA	Q8l4q1 oryza sativa
Q6PB55_MOUSE	Q6pb55 mus musculu
Q4NN47_9DELT	Q4nn47 anaeromyxob
ATG9_PICPA	Q876n4 pichia pasp
Q8NAZ1_HUMAN	Q8naz1 homo sapie
Q4P3M5_USTMA	Q4p3m5 ustilago ma
Q4YZ18_PLABE	Q4yz18 plasmodium

981 36 44.4 953 2 Q5ASZ0 EMENI Q5asz0 aspergillus
982 36 44.4 953 2 Q8IYN5 HUMAN Q8iyn5 homo sapien
983 36 44.4 967 2 Q5CEN2 CRYHO Q5cen2 cryptospori
984 36 44.4 1014 2 Q9MAR9 ARATH Q9mar9 arabidopsis
985 36 44.4 1015 2 Q7XKG6 ORYSA Q7xkg6 oryza sativ
986 36 44.4 1068 2 Q6BJ02 DEBHA Q6bj02 debaryomyce
987 36 44.4 1086 2 Q75K50 ORYSA Q75k50 oryza sativ
988 36 44.4 1092 2 Q6PV87 9NOCA Q6pv87 rhodococcus
989 36 44.4 1131 2 Q5LFG9 BACFN Q5lfg9 bacteroides
990 36 44.4 1131 2 Q64W93 BACFR Q64w93 bacteroides
991 36 44.4 1134 2 Q7X7G6 ORYSA Q7x7g6 oryza sativ
992 36 44.4 1142 2 Q4FY92 LEIMA Q4fy92 leishmania
993 36 44.4 1158 2 Q9LN51 ARATH Q9ln51 arabidopsis
994 36 44.4 1176 2 Q7Y165 ORYSA Q7y165 oryza sativ
995 36 44.4 1184 2 Q55YR2 CRYNE Q55yr2 cryptococcu
996 36 44.4 1184 2 Q5KN33 CRYNE Q5kn33 cryptococcu
997 36 44.4 1190 2 Q9AYJ9 ORYSA Q9ayj9 oryza sativ
998 36 44.4 1191 2 Q9VUY8 DROME Q9vuy8 drosophila
999 36 44.4 1199 2 Q51S54 MAGGR Q51s54 magnaporthe
1000 36 44.4 1204 2 Q6I5B6 ORYSA Q6i5b6 oryza sativ

ALIGNMENTS

RESULT 1
Q6GNG6_XENLA
ID Q6GNG6_XENLA PRELIMINARY; PRT; 508 AA.
AC Q6GNG6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC82822 protein.
GN Name=MGC82822;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC073546; AAH73546.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor.
SQ SEQUENCE 508 AA; 57858 MW; E8019B05639E93FE CRC64;

Query Match 61.7%; Score 50; DB 2; Length 508;
Best Local Similarity 46.2%; Pred. No. 7.9;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTG 14
|::|::|::|::|
Db 214 EVIGPTMLEWLPC 226

RESULT 2
Q4V789_XENTR
ID Q4V789_XENTR PRELIMINARY; PRT; 526 AA.
AC Q4V789;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC098078; AAH98078.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR000219; RhoGEF.

DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS0010; DH 2; 1.
DR PROSITE; PS0003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor; Hypothetical protein.
SQ SEQUENCE 526 AA; 59250 MW; 1E6A5ADAF26C109F CRC64;

Query Match 61.7%; Score 50; DB 2; Length 526;
Best Local Similarity 46.2%; Pred. No. 8.2;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTC 14
|:|:|:|:|
Db 232 EVIGPTMLEWLC 244

RESULT 3
RNS11 HUMAN STANDARD; PRT; 199 AA.
AC Q8TAAI;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable ribonuclease 11 precursor (EC 3.1.27.-) (RNase 11).
GN Name=RNASE11; Synonyms=Cl4orf6; ORFNames=UNQ5832/PRO19669;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
TISSUE=Placenta;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
[2]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).

RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP MEDLINE=2287296; PubMed=12975309; DOI=10.1101/gr.1293003;
RX Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270 (2003).
[4]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; BX161458; CAD61920.1; -; mRNA.
CC EMBL; AK098687; BAC05378.1; -; mRNA.
CC EMBL; AY358794; AAQ89154.1; -; mRNA.
CC EMBL; BC025410; AAH25410.1; -; mRNA.
CC Ensembl; ENSG00000173464; Homo sapiens.
CC HGNC; HGNC:19269; RNASEL1.
CC InterPro; IPR001427; RNaseA.
CC ProDom; PD000535; RNaseA; 1.
KW Endonuclease; Hydrolase; Nuclease; Signal.
FT SIGNAL 1 16 Potential.
FT CHAIN 17 199 Probable ribonuclease 11.
FT ACT_SITE 82 82 Proton acceptor (By similarity).
FT SITE 115 119 Substrate binding (By similarity).
FT DISULFID 98 158 By similarity.
FT DISULFID 114 169 By similarity.
SQ SEQUENCE 199 AA; 22427 MW; BA9DBA907292F1F0 CRC64;

Query Match 60.5%; Score 49; DB 1; Length 199;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | |
Db 184 LTGHSLMSWLVC 195

RESULT 4
QSGAN5 HUMAN
ID QSGAN5_HUMAN PRELIMINARY; PRT; 199 AA.
AC QSGANS;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Ribonuclease 11.
GN Name=RNASE11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15676279; DOI=10.1016/j.ygeno.2004.10.008;
RA Cho S., Beintema J.J., Zhang J.;
RT "The ribonuclease A superfamily of mammals and birds: identifying new
RT members and tracing evolutionary histories.";
RL Genomics 85:208-220(2005).
DR EMBL; AY665806; AAV87184.1; -; mRNA.
SQ SEQUENCE 199 AA; 22427 MW; BA9DBA907292F1F0 CRC64;

Query Match 60.5%; Score 49; DB 2; Length 199;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
Db 184 LTGHSLSMSWLVC 195

RESULT 5
Q9VKC2 DROME
ID Q9VKC2_DROME PRELIMINARY; PRT; 500 AA.
AC Q9VKC2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG12317-PA, isoform A (Cg12317-pb, isoform b) (Amino acid transporter
DE protein JHI-21) (LD39658p).
GN Name=JHI-21; ORFNames=CG12317;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a Genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22418601; PubMed=12530223; DOI=10.1016/S0965-1748(02)00076-0;
RA Dubrovsky E.B., Dubrovskaya V.A., Berger E.M.;
RT "Juvenile hormone signaling during oogenesis in Drosophila
RT melanogaster.";
RL Insect Biochem. Mol. Biol. 32:1555-1565(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003634; AAF53154.1; -; Genomic_DNA.
DR EMBL; AF273478; AAK58692.1; -; mRNA.
DR EMBL; AY051907; AAK93331.1; -; mRNA.
DR Ensembl; CG12317; Drosophila melanogaster.
DR FlyBase; FBgn0028425; CG12317.
DR FlyBase; FBgn0028425; Jhi-21.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR004841; Permease_region.
DR Pfam; PF00324; AA_permease; 1.
KW Transmembrane; Transport.
SQ SEQUENCE 500 AA; 54820 MW; 48BB44291114FEB8 CRC64;

Query Match 60.5%; Score 49; DB 2; Length 500;
Best Local Similarity 69.2%; Pred. No. 12;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTLC 14
Db 71 ESVGSSLLIWLTC 83

RESULT 6
Q7XLY1 ORYSA
ID Q7XLY1_ORYSA PRELIMINARY; PRT; 1510 AA.
AC Q7XLY1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE OSJNBA0042115.9 protein.
GN Name=OSJNBA0042115.9;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL731641; CAE04887.3; -; Genomic DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 1510 AA; 171089 MW; A9E138724D4ACD39 CRC64;

Query Match 60.5%; Score 49; DB 2; Length 1510;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CELVGPSLMSW 11
Db 1370 CQFLGPSLVSW 1380

RESULT 7
Q7XFAQ ORYSA
ID Q7XFAQ_ORYSA PRELIMINARY; PRT; 167 AA.
AC Q7XFAQ;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=OSJNBb0008A05.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017076; AAP53011.1; -; Genomic_DNA.
DR Gramene; Q7XFAQ6; -.
KW Hypothetical protein.
SQ SEQUENCE 167 AA; 18151 MW; A217872B29BFBFB CRC64;

Query Match 58.0%; Score 47; DB 2; Length 167;
Best Local Similarity 61.5%; Pred. No. 8.6;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTLC 14
Db 103 EAAEPSMASWLTLC 115

RESULT 8
Q8W5F2 ORYSA
ID Q8W5F2_ORYSA PRELIMINARY; PRT; 167 AA.
AC Q8W5F2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNBb0008A05.14.
GN Name=OSJNBb0008A05.14;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091749; AAL31079.1; -; Genomic_DNA.
DR Gramene; Q8W5F2; -.
KW Hypothetical protein.

SQ SEQUENCE 167 AA; 18151 MW; A217872B29BFBFB CRC64;

Query Match 58.0%; Score 47; DB 2; Length 167;
Best Local Similarity 61.5%; Pred. No. 8.6;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWLT 14
| :||:|||||
Db 103 EAAEPSMASWLT 115

RESULT 9
Q9HE61_NEUCR
ID Q9HE61_NEUCR PRELIMINARY; PRT; 745 AA.
AC Q9HE61;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Related to MKT1 protein.
GN Name=B2F7.010;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL451013; CAC18153.1; -; Genomic_DNA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR006084; XPGC Rad.
DR PRINTS; PR00853; XPGRADSUPER.
SQ SEQUENCE 745 AA; 84185 MW; 45F372292C354970 CRC64;

Query Match 58.0%; Score 47; DB 2; Length 745;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLT 13
|:||||:|
Db 338 LIGPRLLSWIT 348

RESULT 10
O54512_YEREN
ID O54512_YEREN PRELIMINARY; PRT; 365 AA.
AC O54512;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Irp3 protein.
GN Name=irp3;
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WA-314;
RX MEDLINE=98117033; PubMed=9457855;
RA Pelludat C., Rakin A., Jacobi C., Schubert S., Heesemann J.;
RT "The yersiniabactin biosynthetic gene cluster of Yersinia
enterocolitica: organization and siderophore-dependent regulation.";
RL J. Bacteriol. 180:538-546(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WA-314;

RA Cosima P.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12527; CAA73128.1; -; Genomic_DNA.
DR PIR; T30343; T30343.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR InterPro; IPR010091; Thiaz_red.
DR Pfam; PF01408; GFO_IDH_MocA; 1.
DR TIGRFAMS; TIGR01761; thiaz-red; 1.
SQ SEQUENCE 365 AA; 40705 MW; E0CBB6EC56F3A7D1 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWL 12
|:||||:|
Db 290 CETVGPEGVSWL 301

RESULT 11
Q9Z3C6_YERPE
ID Q9Z3C6_YERPE PRELIMINARY; PRT; 365 AA.
AC Q9Z3C6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE YbtU protein.
GN Name=ybtU;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=6/69;
RA Buchrieser C., Rusniok C., Couve E., Frangeul L., Billault A.,
RA Kunst F., Carniel E., Glaser P.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM6+;
RX MEDLINE=99035519; PubMed=9818149; DOI=10.1016/S1074-5521(98)90115-6;
RA Gehring A.M., DeMoll E., Fetherston J.D., Mori I., Mayhew G.F.,
RA Blattner F.R., Walsh C.T., Perry R.D.;
RT "Iron acquisition in plague: modular logic in enzymatic biogenesis of
yersiniabactin by Yersinia pestis.";
RL Chem. Biol. 5:573-586(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=10231486;
RA Fetherston J.D., Bertolino V.J., Perry R.D.;
RT "YbtP and YbtQ: two ABC transporters required for iron uptake in
Yersinia pestis.";
RL Mol. Microbiol. 32:289-299(1999).
DR EMBL; AL031866; CAA21392.1; -; Genomic_DNA.
DR EMBL; AF091251; AAC69589.1; -; Genomic_DNA.
DR PIR; T17441; T17441.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR InterPro; IPR010091; Thiaz_red.
DR Pfam; PF01408; GFO_IDH_MocA; 1.
DR TIGRFAMS; TIGR01761; thiaz-red; 1.
SQ SEQUENCE 365 AA; 40782 MW; 3DA7E8320682B90E CRC64;

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWL 12
|:||||:|
Db 290 CETVGPEGVSWL 301

RA Menendez C., Igloi G., Henninger H., Brandsch R.;

RT "A pAO1-encoded molybdopterin cofactor gene (moaA) of Arthrobacter

RT nicotinovorans: characterization and site-directed mutagenesis of the

RT encoded protein.";

RL Arch. Microbiol. 164:142-151(1995).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=9808892; PubMed=9428706;

RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,

RA Boettcher B., Brandsch R.;

RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a

RT bacterial plasmid: characterization of MoaA as a filament-forming

RT protein with adenosinetriphosphatase activity.";

RL Eur. J. Biochem. 250:524-531(1997).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=99096870; PubMed=9878353; DOI=10.1006/jmbi.1998.2227;

RA Schenk S., Hoelz A., Kraus B., Decker K.;

RT "Gene structure and properties of enzymes of the plasmid-encoded

RT nicotine catabolism of Arthrobacter nicotinovorans.";

RL J. Mol. Biol. 284:1323-1339(1998).

RN [5]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=97230479; PubMed=9073580; DOI=10.1006/plas.1996.1272;

RA Menendez C., Igloi G.L., Brandsch R.;

RT "IS1473, a putative insertion sequence identified in the plasmid pAO1

RT from Arthrobacter nicotinovorans: isolation, characterisation and

RT distribution among Arthrobacter species.";

RL Plasmid 37:35-41(1997).

RN [6]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=21405725; PubMed=11514508;

RX DOI=10.1128/JB.183.18.5262-5267.2001;

RA Baitsch D., Sandu C., Brandsch R., Igloi G.L.;

RT "A gene cluster on pAO1 of Arthrobacter nicotinovorans involved in the

RT degradation of the plant alkaloid nicotine: cloning, purification and

RT characterization of 2,6-dihydroxypyridine 3-hydroxylase.";

RL J. Bacteriol. 183:5262-5267(2001).

RN [7]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22505657; PubMed=12618462;

RX DOI=10.1128/JB.185.6.1976-1986.2003;

RA Igloi G.L., Brandsch R.;

RT "Sequence of the 165-kilobase catabolic plasmid pAO1 from Arthrobacter

RT nicotinovorans and identification of a pAO1-dependent nicotine uptake

RT system.";

RL J. Bacteriol. 185:1976-1986(2003).

DR EMBL; AJ507836; CAD47969.1; -; Genomic_DNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.

DR GO; GO:0006865; F:amino acid transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR002293; AA/rel_permeasel.

DR InterPro; IPR004841; Permease_region.

DR Pfam; PF00324; AA_permease; 1.

KW Plasmid; Transmembrane; Transport.

SQ SEQUENCE 497 AA; 51904 MW; ED8804A8133B7B44 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 497;

Best Local Similarity 63.6%; Pred. No. 56;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLT 13

Db 381 LIGPALLLWLT 391

RESULT 22

ARHG3_MOUSE

ID ARHG3_MOUSE STANDARD; PRT; 524 AA.

AC Q91X46; Q8CDM0; Q91VY4; Q99K14; Q9DC31;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Rho guanine nucleotide exchange factor 3.

GN Name=Arhgef3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidea; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 3).

RC STRAIN=C57BL/6J; TISSUE=Lung, and Testis;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard R., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).

RC STRAIN=Czech II, and FVB/N; TISSUE=Mammary tumor, and Salivary gland;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP FUNCTION, AND SUBCELLULAR LOCATION.

RX PubMed=11839749; DOI=10.1074/jbc.M11108200;

RA Schmidt A., Hall A.;

RT "The Rho exchange factor Net1 is regulated by nuclear sequestration.";

RL J. Biol. Chem. 277:14581-14588(2002).

CC -!- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for

```
CC CC RhoA and RhoB GTPases.
CC -|- SUBUNIT: Interacts with RHOA and RHOB.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q91X46-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q91X46-2; Sequence=VSP_011613;
CC Name=3;
CC IsoId=Q91X46-3; Sequence=VSP_011614;
CC -|- SIMILARITY: Contains 1 DH (DBL-homology) domain.
CC -|- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AK004600; BAB23401.1; -; mRNA.
CC EMBL; AK029872; BAC26652.1; -; mRNA.
CC EMBL; BC005517; AAH05517.1; -; mRNA.
CC EMBL; BC007153; AAH07153.1; ALT INIT; mRNA.
CC EMBL; BC012262; AAH12262.1; -; mRNA.
CC HSSP; Q15811; 1KI1.
CC Ensembl; ENSMUSG00000021895; Mus musculus.
CC MGI; MGI:1918954; Arhgef3.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH type.
CC InterPro; IPR000219; RhoGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00621; RhoGEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RhoGEF; 1.
CC PROSITE; PS50010; DH 2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC KW Alternative splicing; Guanine-nucleotide releasing factor.
CC DOMAIN 121 303
CC FT DOMAIN 290 448
CC FT VARSPLIC 1 31
CC FT MVAKDYPFYLTVKRANCSLEAPLGSGVAKDE -> MRSERP
CC FT MVWCCFFLRAQKRKQSSQDEDAVSLCSLDIS (in
CC FT isoform 2).
CC FT /FTId=VSP_011613.
CC FT MVAKDYPFYLTVKRANCSLEAPLGSGVAKDE -> MFPSPK
CC FT ACNFRGRKRKQSSQDEDAVSLCSLDIS (in isoform
CC FT 3).
CC FT /FTId=VSP_011614.
CC FT V -> I (in Ref. 2; AAH05517).
CC FT CONFLICT 363 363
CC SEQUENCE 524 AA; 59526 MW; 579FF96C1F0B0ADA CRC64;
Query Match 55.6%; Score 45; DB 1; Length 524;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 ELVGPSLMSWLTC 14
Db 194 EHVGPILVGWLPC 206
RESULT 23
ARHG3_HUMAN STANDARD; PRT; 526 AA.
ID ARHG3_HUMAN Q6NUN3; Q7Z4U2; Q7Z5T2; Q9H7T4;
AC Q9NR81; Q6NUN3; Q7Z4U2; Q7Z5T2; Q9H7T4;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rho guanine nucleotide exchange factor 3 (Exchange factor found in
DE platelets and leukemic and neuronal tissues) (XPLN).
GN Name=ARHGEF3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
```

```
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=20334277; PubMed=10873612; DOI=10.1006/bbrc.2000.2925;
RA Thiesen S., Kuebart S., Ropers H.-H., Nothwang H.G.;
RT "Isolation of two novel human RhoGEFs, ARHGEF3 and ARHGEF4, in 3p13-21
RT and 2q22.";
RL Biochem. Biophys. Res. Commun. 273:364-369 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RX TISSUE=Testis;
RC MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansonge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2), AND VARIANT VAL-335.
RA Guo J.H., Yu L.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBDJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
RP VAL-335.
RC TISSUE=Brain, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 71-526 (ISOFORMS 1/2).
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamiyara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
```


RA Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [6]
RP FUNCTION, AND TISSUE SPECIFICITY.
RX PubMed=12221096; DOI=10.1074/jbc.M207401200;
RA Arthur W.T., Ellierbroek S.M., Der C.J., Burridge K., Wennerberg K.;
RT "XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not
RT RhoC.";
RL J. Biol. Chem. 277:42964-42972(2002).
CC -!- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for
CC RhoA and RhoB GTPases.
CC -!- SUBUNIT: Interacts with RHOA and RHOB.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NR81-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NR81-2; Sequence=VSP_011612;
CC -!- TISSUE SPECIFICITY: Widely expressed. Highest levels are found in
CC adult brain and skeletal muscle. Lower levels are found in heart
CC and kidney.
CC -!- SIMILARITY: Contains 1 DH (DBL-homology) domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF249744; AAF79954.1; -; mRNA.
CC EMBL; AL136832; CAB66766.1; -; mRNA.
CC EMBL; AF433662; AAP97313.1; -; mRNA.
CC EMBL; BC054345; AAH54345.1; ALT_INIT; mRNA.
CC EMBL; BC068513; AAH68513.1; ALT_INIT; mRNA.
CC EMBL; AK024340; BAB14891.1; ALT_INIT; mRNA.
CC HSSP; Q15811; 1KI1.
CC Ensembl; ENSG00000163947; Homo sapiens.
CC HGNC; HGNC:683; ARHGEF3.
CC GO; GO:0005089; F:Rho guanyl-nucleotide exchange factor activity; TAS.
CC GO; GO:0007266; P:Rho protein signal transduction; TAS.
CC InterPro; IPR001331; GDS_CDC24.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH_type.
CC InterPro; IPR000219; RhoGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00621; RhoGEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RhoGEF; 1.
CC SMART; SM00325; RhoGEF; 1.
CC PROSITE; PS00741; DH_1; FALSE_NEG.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
KW Alternative splicing; Guanine-nucleotide releasing factor;
KW Polymorphism.
FT DOMAIN 122 304 DH.
FT DOMAIN 291 449 PH.
FT VARSPLIC 1 32
FT MVADYFFYLTIVKRNCSLELPASGPAKDAE -> MDSST
FT AMNQSCRGMEENKERPKRQRQNNFPMFPSPKAWFRGRK
FT KQSTQDEDAVSLCSLDIS (in isoform 2).
FT /FTid=VSP_011612.
FT K -> R (in dbSNP:3732507).
FT /FTid=VAR_021935.

FT VARIANT 335 335 L -> V (in dbSNP:3772219).
FT /FTid=VAR_021936.
FT CONFLICT 410 410 I -> T (in Ref. 3).
SQ SEQUENCE 526 AA; 59783 MW; 0FE8248AB52D1C8C CRC64;

Query Match 55.6%; Score 45; DB 1; Length 526;
Best Local Similarity 61.5%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELVGPSILMSWLTC 14
Db 195 EHVGPILVGWLPC 207

RESULT 24
ARHG3 MACFA
ID ARHG3 MACFA STANDARD; PRT; 526 AA.
AC Q9N0A8;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rho guanine nucleotide exchange factor 3.
GN Name=ARHGEF3; ORFNames=QcCE-16434;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries".
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for
CC RhoA and RhoB GTPases (By similarity).
CC -!- SUBUNIT: Interacts with RHOA and RHOB (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DH (DBL-homology) domain.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB046022; BAB01604.1; -; mRNA.
CC HSSP; Q15811; 1KI1.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH_type.
CC InterPro; IPR000219; RhoGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00621; RhoGEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RhoGEF; 1.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor.
FT DOMAIN 122 304 DH.
FT DOMAIN 291 449 PH.
SQ SEQUENCE 526 AA; 59764 MW; 0FE83450D52D1C8C CRC64;

Query Match 55.6%; Score 45; DB 1; Length 526;
Best Local Similarity 61.5%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELVGPSILMSWLTC 14
Db 195 EHVGPILVGWLPC 207

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC098272; AAH98272.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhGEF.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 566 AA; 63759 MW; D45197FD63AFAD72 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 566;
Best Local Similarity 61.5%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTC 14
| | | | | : | | |
Db 235 EHVGPILVGWLPC 247

RESULT 29
Q52BN4 MAGGR
ID Q52BN4_MAGGR PRELIMINARY; PRT; 734 AA.
AC Q52BN4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG08303.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.

OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H., Gnerre S.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Hafez N.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Higgins H.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkison J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACU01000328; EAA57334.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 734 AA; 81983 MW; 316892F03309C9A5 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 734;
Best Local Similarity 54.5%; Pred. No. 83;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLT 13
| : | | | : | | : |
Db 339 LIGPRMLSWIT 349

RESULT 30
Q8YZ60 ANASP
ID Q8YZ60_ANASP PRELIMINARY; PRT; 907 AA.

Db 610 VNPSLVSWLT 619

```
RESULT 35
Q5SQ40 HUMAN
ID Q5SQ40_HUMAN PRELIMINARY; PRT; 1029 AA.
AC Q5SQ40;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-013;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18505.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1029 AA; 108672 MW; B694B35E126B3091 CRC64;
```

Query Match 54.3%; Score 44; DB 2; Length 1029;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| |||:||||
Db 900 VNPSLVSWLT 909

```
RESULT 36
Q5SQ42 HUMAN
ID Q5SQ42_HUMAN PRELIMINARY; PRT; 1077 AA.
AC Q5SQ42;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-019;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18503.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1077 AA; 113455 MW; CB94D1C58A1E3D21 CRC64;
```

Query Match 54.3%; Score 44; DB 2; Length 1077;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| |||:||||
Db 900 VNPSLVSWLT 909

```
RESULT 37
Q5SQ43 HUMAN
ID Q5SQ43_HUMAN PRELIMINARY; PRT; 1078 AA.
AC Q5SQ43;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-020;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18502.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1078 AA; 113909 MW; 4C8CB9F9FE02CF79 CRC64;
```

Query Match 54.3%; Score 44; DB 2; Length 1078;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| |||:||||
Db 900 VNPSLVSWLT 909

```
RESULT 38
Q6MG49 RAT
ID Q6MG49_RAT PRELIMINARY; PRT; 1096 AA.
AC Q6MG49;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HLA-B associated transcript 3, rat orthologue.
GN Name=Bat3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brown Norway;
RX PubMed=15060004; DOI=10.1101/gr.1987704;
RA Hurt P., Walter L., Sudbrak R., Klages S., Mueller I., Shiina T.,
RA Inoko H., Lehrach H., Guenther E., Reinhardt R., Himmelbauer H.;
RT "The genomic sequence and comparative analysis of the rat major
RT histocompatibility complex.";
RL Genome Res. 14:631-639(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brown Norway;
RA Boehm S., Borzym K., Gelling S., Gimmel V., Heitmann K., Kosiura A.,
RA Lang N., Lehrack S., Thiel J., Sontag M., Hurt P., Himmelbauer H.,
RA Sudbrak R., Reinhardt R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX883045; CAE83997.1; -; Genomic_DNA.
DR HSSP; Q862M4; 1AAR.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
```

Query Match 54.3%; Score 44; DB 2; Length 1078;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| |||:||||
Db 900 VNPSLVSWLT 909

DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1096 AA; 114646 MW; E372D70E024D8F15 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1096;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 919 VNPSLVSWLT 928

RESULT 39
Q9WTN8 RAT PRELIMINARY; PRT; 1098 AA.
AC Q9WTN8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3.
GN Name=BAT3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99317047; PubMed=10390159; DOI=10.1089/104454999315222;
RA Ozaki T., Hanaoka E., Naka M., Nakagawara A., Sakiyama S.;
RT "Cloning and characterization of rat BAT3 cDNA."
RL DNA Cell Biol. 18:503-512(1999).
DR EMBL; AB018791; BAA76607.1; -; mRNA.
DR HSSP; Q9SHE7; 1BT0.
DR Ensembl; ENSRNOG0000000851; Rattus norvegicus.
DR RGD; 71064; Bat3.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1098 AA; 114931 MW; 5246FFE7D5CD8FB3 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1098;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 921 VNPSLVSWLT 930

RESULT 40
Q5HYL9 HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q5HYL9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein DKFZp686L0653.
GN Name=DKFZp686L0653;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endometrium;
RG The German cDNA Consortium;

RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX647244; CAI46045.1; -; mRNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1126 AA; 118724 MW; 967AEE046DED3ABE CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1126;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 900 VNPSLVSWLT 909

RESULT 41
Q96SA6 HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q96SA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3 protein.
GN Name=BAT3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BA000025; BAB63390.1; -; Genomic_DNA.
DR HSSP; Q9SHE7; 1BT0.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1126 AA; 118702 MW; 8A67290B45B46ABA CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1126;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 900 VNPSLVSWLT 909

RESULT 42
Q9BCN4 HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q9BCN4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE HLA-B associated transcript-3, isoform b.
GN Name=BAT3;

GN ORFNames=DAQB-195H10.3-001, XXbac-BCX270M2.3-001,
GN XXbac-BPG296P20.4-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Tracey A.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Griffiths C.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003133; AAH03133.1; -; mRNA.
DR EMBL; AL662801; CAI18315.1; -; Genomic_DNA.
DR EMBL; AL662847; CAI17658.1; -; Genomic_DNA.
DR EMBL; AL670886; CAI17784.1; -; Genomic_DNA.
DR EMBL; AL805934; CAI18501.1; -; Genomic_DNA.
DR EMBL; AL670886; CAI17658.1; JOINED; Genomic_DNA.
DR EMBL; AL662847; CAI17784.1; JOINED; Genomic_DNA.
DR HSSP; Q9SHE7; 1BT0.
DR Ensembl; ENSG00000096155; Homo sapiens.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1126 AA; 118692 MW; 8A67290BC8176ABA CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1126;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLMSWLT 13
| |||: ||||
Db 900 VNPSLVSWLT 909

RESULT 43
BAT3_HUMAN STANDARD; PRT; 1132 AA.
ID BAT3_HUMAN
AC P46379;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Large proline-rich protein BAT3 (HLA-B-associated transcript 3) (G3).
GN Name=BAT3; Synonyms=G3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=T-cell;
RX MEDLINE=90192810; PubMed=2156268;
RA Banerji J., Sands J., Strominger J.L., Spies T.;
RT "A gene pair from the human major histocompatibility complex encodes
RT large proline-rich proteins with multiple repeated motifs and a single
RT ubiquitin-like domain";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378 (1990).
RN [2]
RP PHOSPHORYLATION SITES SER-964; SER-973 AND SER-1117.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
RA Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135 (2004).
CC -!- FUNCTION: Unknown.
CC -!- INTERACTION:
CC Q86WS7;-; NbExp=1; IntAct=EBI-347552, EBI-372406;
CC Q9NWX1:RNF126; NbExp=1; IntAct=EBI-347552, EBI-347629;
CC Q96S82:SB132; NbExp=1; IntAct=EBI-347552, EBI-348604;
CC O43765:SGTA; NbExp=1; IntAct=EBI-347552, EBI-347996;
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; M33519; AAA35587.1; -; mRNA.
DR EMBL; M33521; AAA35588.1; -; Genomic_DNA.
DR EMBL; M33520; AAA35588.1; JOINED; Genomic_DNA.
DR PIR; A35098; A35098.
DR HSSP; P02248; 1UD7.
DR IntAct; P46379; -.
DR Ensembl; ENSG00000096155; Homo sapiens.
DR HGNC; HGNC:13919; BAT3.
DR H-InvDB; HIX0005722; -.
DR MIM; 142590; -.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Phosphorylation; Repeat.
FT DOMAIN 17 77 Ubiquitin-like.
FT REPEAT 242 270 1.
FT REPEAT 415 443 2.
FT REPEAT 574 602 3.
FT REPEAT 608 636 4.
FT REGION 242 636 4 X 29 AA approximate repeats.
FT COMPIAS 202 207 Poly-Pro.
FT COMPIAS 657 670 Poly-Pro.
FT MOD_RES 964 964 Phosphoserine.
FT MOD_RES 973 973 Phosphoserine.
FT MOD_RES 1117 1117 Phosphoserine.
SQ SEQUENCE 1132 AA; 119504 MW; E28CA8A78C38DD18 CRC64;

Query Match 54.3%; Score 44; DB 1; Length 1132;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
Db 906 VNPSLVSWLT 915

RESULT 44
Q5STC1_HUMAN PRELIMINARY; PRT; 1132 AA.
AC Q5STC1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3;
GN ORFNames=DAQB-195H10.3-002, XXbac-BCX270M2.3-002,
GN XXbac-BPG296P20.4-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tracey A.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Griffiths C.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL662847; CAI17659.1; -; Genomic_DNA.
DR EMBL; AL670886; CAI17785.1; -; Genomic_DNA.
DR EMBL; AL805934; CAI18504.1; -; Genomic_DNA.
DR EMBL; AL662801; CAI18314.1; -; Genomic_DNA.
DR EMBL; AL670886; CAI17659.1; JOINED; Genomic_DNA.
DR EMBL; AL662847; CAI17785.1; JOINED; Genomic_DNA.
DR Ensembl; ENSG00000096155; Homo sapiens.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1132 AA; 119409 MW; 625B5F86321367ED CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1132;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
Db 906 VNPSLVSWLT 915

RESULT 45
Q9Z1R2_MOUSE PRELIMINARY; PRT; 1154 AA.
AC Q9Z1R2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3.
GN Name=Bat3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129;
RX PubMed=14656967; DOI=10.1101/gr.1736803;
RA Xie T., Rowen L., Aguado B., Ahearn M.E., Madan A., Qin S.,
RA Campbell R.D., Hood L.;
RT "Analysis of the gene-dense major histocompatibility complex class III
region and its comparison to mouse."
RL Genome Res. 13:2621-2636(2003).
DR EMBL; AF109719; AAC82479.1; -; Genomic_DNA.
DR HSSP; Q9SHE7; 1BT0.
DR Ensembl; ENSMUSG00000024392; Mus musculus.
DR MGI; MGI:1919439; Bat3.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1154 AA; 121037 MW; 7F3FD14DF5AC1211 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1154;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
Db 928 VNPSLVSWLT 937

RESULT 46
Q5SQ39_HUMAN PRELIMINARY; PRT; 1162 AA.
AC Q5SQ39;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-018;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18506.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1162 AA; 122341 MW; FE67886ACE6F6432 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1162;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
Db 936 VNPSLVSWLT 945

RESULT 47
O95874_HUMAN PRELIMINARY; PRT; 1229 AA.
ID O95874_HUMAN
AC O95874;


```
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656967; DOI=10.1101/gr.1736803;
RA Xie T., Rowen L., Aguado B., Ahearn M.E., Madan A., Qin S.,
RA Campbell R.D., Hood L.;
RT "Analysis of the gene-dense major histocompatibility complex class III
RT region and its comparison to mouse.";
RL Genome Res. 13:2621-2636(2003).
DR EMBL; AF129756; AAD18085.1; -; Genomic_DNA.
DR HSSP; Q9SHE7; 1BT0.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1229 AA; 130200 MW; 2EA1F2AFB6DCD221 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 1003 VNPSLVSWLT 1012

RESULT 48
Q5SQ35 HUMAN
ID Q5SQ35_HUMAN PRELIMINARY; PRT; 1229 AA.
AC Q5SQ35;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3 (Fragment).
GN Name=BAT3; ORFNames=DAQB-195H10.3-014;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18510.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
FT NON_TER 1 1
SQ SEQUENCE 1229 AA; 130200 MW; 2EA1F2AFB6DCD221 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
| | | | |
Db 1003 VNPSLVSWLT 1012
```

```
RESULT 49
Q4HCM7_9DE10
ID Q4HCM7_9DE10 PRELIMINARY; PRT; 108 AA.
AC Q4HCM7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DgeODRAFT_2030;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israeli S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000001; EAL83924.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11647 MW; 76EC53C1650B1858 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 108;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 GPSLMSWLT 14
| | | | |
Db 48 GPAVSGWLT 57

RESULT 50
Q9Z537_STRCO
ID Q9Z537_STRCO PRELIMINARY; PRT; 224 AA.
AC Q9Z537;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SCO6135; ORFNames=SC9B2.22c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
```

RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939126; CAA22802.1; -; Genomic_DNA.
DR FIR; T35918; T35918.
KW Complete proteome.
SQ SEQUENCE 224 AA; 23633 MW; ACFAAA6D2F502F8C CRC64;

Query Match 53.1%; Score 43; DB 2; Length 224;
Best Local Similarity 72.7%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWL 12
| | | | | | | |
Db 162 ELVGPELWPWL 172

Search completed: May 12, 2006, 10:50:20
Job time : 134.077 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 21.8462 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-80
Perfect score: 97
Sequence: 1 TIKGPTLRQWLKSRHTS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	50.5	122	2 AI0380	conserved hypothet
2	47	48.5	327	2 B71900	hypothetical prote
3	45	46.4	186	2 A90167	adenylate cyclase,
4	44	45.4	161	2 T06826	beta-fructofuranos
5	44	45.4	1019	2 T11560	pol polyprotein -
6	44	45.4	1058	2 S08436	pol polyprotein -
7	43	44.3	63	2 T30614	hypothetical prote
8	43	44.3	216	2 S35151	photosystem I chai
9	43	44.3	217	2 S46354	pol polyprotein -
10	43	44.3	233	2 A83862	initiation of chro
11	43	44.3	241	2 S07740	hypothetical prote
12	43	44.3	269	2 S73999	hypothetical prote
13	43	44.3	656	2 S30484	pol polyprotein -
14	43	44.3	656	2 S30483	pol polyprotein -
15	43	44.3	1583	2 S59644	sister chromatid c
16	42.5	43.8	275	2 AC0189	probable exported
17	42	43.3	114	2 E71171	hypothetical prote
18	42	43.3	171	2 T20567	hypothetical prote
19	42	43.3	416	2 D72456	probable glutamyl-
20	42	43.3	438	2 S71157	cytochrome c bioge
21	42	43.3	519	2 C86160	hypothetical prote
22	42	43.3	581	2 T12095	beta-fructofuranos
23	42	43.3	617	2 S75447	proline-tRNA ligas
24	42	43.3	1061	1 GNLJG4	HIV-1 retropepsin
25	42	43.3	2108	2 H70819	probable polyketid
26	41.5	42.8	877	2 T03098	p97 protein - Toxo
27	41	42.3	70	2 T06920	ribosomal protein
28	41	42.3	153	2 A97524	hypothetical prote
29	41	42.3	178	2 AB2743	hypothetical prote

41	42.3	188	2	C82863	recombinase XFa001
41	42.3	203	2	T40206	hypothetical prote
41	42.3	238	2	S76936	hypothetical prote
41	42.3	257	2	E70429	tRNA guanine-N1 me
41	42.3	264	2	AG2095	hypothetical prote
41	42.3	306	2	D70601	UTP-glucose-1-phos
41	42.3	380	2	B47029	methylease Llapi -
41	42.3	397	2	G69449	tryptophan synthas
41	42.3	419	1	S29127	carboxypeptidase A
41	42.3	472	2	F70876	probable papA3 pro
41	42.3	622	2	S35122	site-specific DNA-
41	42.3	1039	2	S46347	pol polyprotein -
41	42.3	1123	2	T51517	telomerase reverse
41	42.3	1140	2	S73786	hypothetical prote
41	42.3	1191	2	T31091	hypothetical prote
41	42.3	3345	2	T13423	hypothetical prote
40	41.2	158	2	D72305	hypothetical prote
40	41.2	185	2	T49611	hypothetical prote
40	41.2	200	2	T23485	hypothetical prote
40	41.2	207	2	T37464	probable glutathio
40	41.2	236	1	T24395	dihydropteridine r
40	41.2	242	2	C70895	hypothetical prote
40	41.2	249	2	E87575	ABC transporter, A
40	41.2	262	2	B83126	probable transcrip
40	41.2	306	2	T45453	UTP-glucose-1-phos
40	41.2	336	2	E72389	hypothetical prote
40	41.2	344	2	E84377	protein export [im
40	41.2	349	2	B97912	dtDPglucose 4,6-de
40	41.2	395	2	S40171	phosphoprotein pho
40	41.2	419	1	CPRTA	carboxypeptidase A
40	41.2	505	2	T19971	hypothetical prote
40	41.2	506	2	T19973	hypothetical prote
40	41.2	521	2	T01923	hypothetical prote
40	41.2	527	2	B64633	hypothetical prote
40	41.2	612	2	JQ1346	glucan 1,4-alpha-g
40	41.2	978	2	B89971	conserved hypothet
40	41.2	1009	2	S28081	pol polyprotein -
40	41.2	1009	2	S44621	C50C3.2 protein -
40	41.2	1034	1	GNLJCA	HIV-1 retropepsin
40	41.2	1035	1	GNLJGG	HIV-1 retropepsin
40	41.2	1036	1	GNLJG2	HIV-1 retropepsin
40	41.2	1040	2	T08190	hypothetical prote
40	41.2	1054	1	GNLJG5	HIV-1 retropepsin
40	41.2	1055	1	GNLJST	HIV-1 retropepsin
40	41.2	1055	2	S53092	pol polyprotein -
40	41.2	1056	1	GNLJG3	HIV-1 retropepsin
40	41.2	1712	1	CGHU2B	collagen alpha 2(I
40	41.2	2609	2	T40399	probable transport
39.5	40.7	642	2	A83268	probable soluble l
39.5	40.7	722	2	T37970	probable G2-specif
39.5	40.7	1758	2	T34393	hypothetical prote
39	40.2	195	2	T36141	probable nicotinam
39	40.2	220	2	AC0318	probable nicotinam
39	40.2	240	2	S73922	uracil DNA glycosyl
39	40.2	240	2	AE1145	phosphoribosylform
39	40.2	314	2	H90638	hypothetical prote
39	40.2	318	2	S17197	nitrate reductase
39	40.2	320	2	H85489	hypothetical prote
39	40.2	327	2	E82277	probable transposa
39	40.2	331	2	B48445	glyceraldehyde-3-p
39	40.2	336	2	A47542	short-chain alcoho
39	40.2	349	2	F86649	dtDP-glucose 4,6-d
39	40.2	373	2	D64729	probable activator
39	40.2	410	2	H86290	hypothetical prote
39	40.2	533	2	G83658	hypothetical prote
39	40.2	567	2	S69778	adhesin AP65-1 pre
39	40.2	571	2	S48384	DNA43 protein - ye
39	40.2	615	2	T15575	hypothetical prote
39	40.2	644	1	W1WL58	E1 protein - human
39	40.2	721	2	A39707	erythrocyte membra
39	40.2	721	2	A33319	microtubule-associ
39	40.2	859	2	G86242	hypothetical prote
39	40.2	877	2	S72541	nitrate reductase

103 39 40.2 989 2 T15576 hypothetical prote
104 39 40.2 1055 1 A37205 leukotoxin A - Act
105 39 40.2 1339 2 T40245 probable transcrip
106 39 40.2 1544 2 E59431 phosphoinositide-b
107 38.5 39.7 256 2 D90026 hypothetical prote
108 38.5 39.7 416 1 JC4952 transcription init
109 38.5 39.7 481 2 T49321 related to heterok
110 38.5 39.7 529 2 AI0986 probable membrane
111 38.5 39.7 1073 2 I51055 recombination acti
112 38.5 39.7 2476 2 T34022 zonadhesin - pig
113 38 39.2 104 2 A96002 hypothetical prote
114 38 39.2 134 2 B75468 hypothetical prote
115 38 39.2 182 2 H95917 hypothetical prote
116 38 39.2 194 2 S70990 hypothetical prote
117 38 39.2 241 2 T23428 hypothetical prote
118 38 39.2 250 2 T14548 beta-fructofuranos
119 38 39.2 254 1 G69592 branched-chain ami
120 38 39.2 256 1 MFNZ matrix protein - h
121 38 39.2 257 2 F72116 conserved hypothet
122 38 39.2 257 2 E86507 hypothetical prote
123 38 39.2 260 2 S72748 B1177 F3 136 prote
124 38 39.2 281 2 T04858 hypothetical prote
125 38 39.2 311 1 RGECK regulatory protein
126 38 39.2 311 2 AH0867 transcriptions acti
127 38 39.2 311 2 C85936 positive regulator
128 38 39.2 311 2 H91090 positive regulator
129 38 39.2 321 2 T42591 gene 48 protein -
130 38 39.2 327 2 A64613 conserved hypothet
131 38 39.2 331 2 A72514 hypothetical prote
132 38 39.2 335 2 E89819 hypothetical prote
133 38 39.2 342 1 S64042 porphobilinogen sy
134 38 39.2 346 2 T10173 sterol 24-C-methyl
135 38 39.2 356 2 C87699 dTDP-D-glucose-4,6
136 38 39.2 357 2 C89880 hypothetical prote
137 38 39.2 367 2 T06780 probable sterol 24
138 38 39.2 368 2 G65119 hypothetical 40.4
139 38 39.2 368 2 H95203 GTP-binding protei
140 38 39.2 368 2 H98070 conserved hypothet
141 38 39.2 375 2 A47117 formate dehydrogen
142 38 39.2 393 2 B85992 probable transport
143 38 39.2 399 2 F91146 probable transport
144 38 39.2 438 2 E97342 high affinity gluc
145 38 39.2 473 2 E84853 hypothetical prote
146 38 39.2 533 2 S43526 amidophosphoribosy
147 38 39.2 610 2 S19461 probable membrane
148 38 39.2 623 2 D71435 hypothetical prote
149 38 39.2 632 2 AI1607 D-1-deoxyxylulose
150 38 39.2 636 2 AI2552 transposase alr801
151 38 39.2 649 2 S74823 N-acetylmuramoyl-L
152 38 39.2 670 2 T02092 beta-fructofuranos
153 38 39.2 735 2 T41187 hypothetical prote
154 38 39.2 807 2 T24110 hypothetical prote
155 38 39.2 952 2 E86147 tIN6.4 protein - A
156 38 39.2 1004 2 T30641 probable DNA-direc
157 38 39.2 1045 2 S23570 pol polyprotein ho
158 38 39.2 1102 2 S55100 hypothetical prote
159 38 39.2 1114 2 B86423 hypothetical prote
160 38 39.2 1123 2 T28139 PK4 protein kinase
161 38 39.2 1124 1 GNLFJP HIV-1 retropepsin
162 38 39.2 1124 2 S23822 pol polyprotein -
163 38 39.2 1499 2 A89813 glutamate synthase
164 38 39.2 1522 2 C96578 hypothetical prote
165 38 39.2 2078 2 T25400 hypothetical prote
166 38 39.2 2285 1 G02434 DNA-directed DNA p
167 37.5 38.7 142 2 B82223 hypothetical prote
168 37.5 38.7 291 2 H81179 membrane protein N
169 37.5 38.7 291 2 F81923 probable ABC-trans
170 37.5 38.7 296 2 AG0147 probable membrane
171 37.5 38.7 513 1 S50216 translation initia
172 37.5 38.7 753 2 AH0097 probable biotin su
173 37 38.1 77 2 A83468 hypothetical prote
174 37 38.1 87 2 S42171 cytochrome-c oxida
175 37 38.1 139 2 S77504 ycf21 protein - Sy

176 37 38.1 180 2 T09063
177 37 38.1 209 2 A83849
178 37 38.1 214 2 AI2371
179 37 38.1 232 2 S40165
180 37 38.1 241 2 G90190
181 37 38.1 255 2 AE0537
182 37 38.1 265 2 A56838
183 37 38.1 268 2 D33465
184 37 38.1 273 2 AB0141
185 37 38.1 276 2 D82169
186 37 38.1 319 2 S59407
187 37 38.1 326 2 C24430
188 37 38.1 327 2 H82736
189 37 38.1 330 2 C87036
190 37 38.1 336 1 DEP2G
191 37 38.1 337 2 A35080
192 37 38.1 338 1 DEIS3C
193 37 38.1 338 2 JQ1287
194 37 38.1 339 2 AG0799
195 37 38.1 340 2 AD0701
196 37 38.1 352 2 B83000
197 37 38.1 361 2 AF0767
198 37 38.1 361 2 S78542
199 37 38.1 361 2 H64969
200 37 38.1 361 2 S15299
201 37 38.1 361 2 S41534
202 37 38.1 368 2 F90152
203 37 38.1 370 2 AB2578
204 37 38.1 370 2 A97360
205 37 38.1 373 2 AB0604
206 37 38.1 379 2 JQ2272
207 37 38.1 396 2 AF2835
208 37 38.1 407 2 A86298
209 37 38.1 409 2 T47118
210 37 38.1 413 2 A97613
211 37 38.1 420 2 C83246
212 37 38.1 422 2 F96826
213 37 38.1 428 2 B71403
214 37 38.1 433 2 S51837
215 37 38.1 433 2 S51836
216 37 38.1 437 2 T45481
217 37 38.1 469 2 A99656
218 37 38.1 490 2 T38088
219 37 38.1 491 2 G89860
220 37 38.1 499 2 B85507
221 37 38.1 502 2 S43123
222 37 38.1 525 2 A85362
223 37 38.1 525 2 C69794
224 37 38.1 536 1 SYECBB
225 37 38.1 536 2 E85558
226 37 38.1 536 2 A99708
227 37 38.1 540 2 AC0063
228 37 38.1 547 2 T48551
229 37 38.1 567 2 S69779
230 37 38.1 568 2 E90364
231 37 38.1 580 2 C82551
232 37 38.1 586 2 B84271
233 37 38.1 631 2 B87250
234 37 38.1 664 2 G89894
235 37 38.1 752 2 T35244
236 37 38.1 775 2 AB3402
237 37 38.1 816 2 A71006
238 37 38.1 860 2 C86203
239 37 38.1 907 2 AB1885
240 37 38.1 984 2 T48216
241 37 38.1 990 1 G46335
242 37 38.1 1032 2 S12153
243 37 38.1 1204 2 T18812
244 37 38.1 1254 2 A54818
245 37 38.1 1333 2 A37488
246 37 38.1 1336 2 S25716
247 37 38.1 1420 2 A32869
248 37 38.1 1514 2 T34869

hypothetical prote
alpha-ribazole-5'-
hypothetical prote
glutathione transf
conserved hypothet
hypothetical prote
uroporphyrinogen-I
lic-1 protein D -
conserved hypothet
conserved hypothet
ribosomal protein
glyceraldehyde-3-p
hypothetical prote
probable conserved
glyceraldehyde-3-p
glyceraldehyde-3-p
glyceraldehyde-3-p
glyceraldehyde-3-p
probable transcrip
tetrathionate redu
dTDP-D-glucose 4,6
dTDP-glucose 4,6-d
dTDPglucose 4,6-de
dTDPglucose 4,6-de
dTDPglucose 4,6-de
dTDPglucose 4,6-de
sun (fmu) protein
membrane-bound lyt
outer membrane lip
probable oxidoredu
formate dehydrogen
sarcosine oxidase
hypothetical prote
thiamine pyridinyl
sarcosine oxidase
probable binding p
hypothetical prote
hypothetical prote
glyceraldehyde-3-p
glyceraldehyde-3-p
O-acetylhomoserine
hypothetical prote
protoporphyrinogen
hypothetical prote
hypothetical prote
protein-tyrosine-p
leucyl aminopeptid
glutamate synthase
2,3-dihydroxybenzo
2,3-dihydroxybenzo
2,3-dihydroxybenzo
hypothetical prote
fructosidase-like
adhesin AP65-2 pre
hypothetical prote
phage-related term
glutamyl-tRNA synt
dnaK protein [impo
protein kinase [im
UvrA-like ABC tran
topoisomerase IV c
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
env polyprotein pr
pol polyprotein -
hypothetical prote
myosin-VI [similar
Ras guanine nucleo
Ras guanine nucleo
apolipoprotein(a)
glutamate synthase

249 37 38.1 2183 1 G48556 genome polyprotein 322
250 37 38.1 2183 1 ZLNZMV genome polyprotein 323
251 37 38.1 2319 2 A47004 coagulation factor 324
252 37 38.1 2329 2 S44625 C50C3.6 protein - 325
253 37 38.1 2363 2 T38841 probable pre-mRNA 326
254 37 38.1 2427 2 T16613 hypothetical prote 327
255 37 38.1 2605 2 T18552 saframycin MxI syn 328
256 37 38.1 3430 1 GNWVWV genome polyprotein 329
257 37 38.1 3433 1 GNWVKV genome polyprotein 330
258 37 38.1 4485 2 T08044 dynein gamma heavy 331
259 36.5 37.6 172 2 T36107 probable serine/ar 332
260 36.5 37.6 341 2 C70932 probable dehydroge 333
261 36.5 37.6 515 1 A59309 interferon-inducib 334
262 36.5 37.6 746 2 F95890 probable aldehyde 335
263 36.5 37.6 830 2 T37973 rad16 nucleotide e 336
264 36.5 37.6 870 2 A96637 hypothetical prote 337
265 36.5 37.6 1876 2 T28627 vitellogenin - Rip 338
266 36 37.1 64 2 S41873 transcription fact 339
267 36 37.1 86 2 AE2441 hypothetical prote 340
268 36 37.1 107 2 S52710 probable neuroept 341
269 36 37.1 111 2 F75052 hypothetical prote 342
270 36 37.1 114 2 E86353 protein F2E2.12 [i 343
271 36 37.1 123 1 LAHO alpha-lactalbumin 344
272 36 37.1 123 2 S28933 alpha-lactalbumin 345
273 36 37.1 124 2 F72747 hypothetical prote 346
274 36 37.1 130 2 I83571 probable membrane 347
275 36 37.1 130 2 C85701 hypothetical prote 348
276 36 37.1 130 2 F90843 hypothetical prote 349
277 36 37.1 131 2 S74539 hypothetical prote 350
278 36 37.1 151 1 S45108 hypothetical prote 351
279 36 37.1 160 2 C70947 hypothetical prote 352
280 36 37.1 169 2 S26011 hypothetical prote 353
281 36 37.1 172 2 AD3606 molybdopterin bios 354
282 36 37.1 188 2 B25317 gag polyprotein - 355
283 36 37.1 203 2 H96525 probable terpene c 356
284 36 37.1 218 2 JC7220 nuclear protein SR 357
285 36 37.1 219 2 F70574 hypothetical prote 358
286 36 37.1 219 2 B83889 hypothetical prote 359
287 36 37.1 221 2 E64305 conserved hypothet 360
288 36 37.1 229 2 JC7219 nuclear protein SR 361
289 36 37.1 231 2 S76204 hypothetical prote 362
290 36 37.1 239 2 A25317 gag polyprotein - 363
291 36 37.1 252 2 S06567 finger protein (cl 364
292 36 37.1 254 2 S35743 p19 protein - avia 365
293 36 37.1 259 2 AB3572 succinoglycan bios 366
294 36 37.1 267 2 I40327 baf protein - Bord 367
295 36 37.1 279 2 E83456 hypothetical prote 368
296 36 37.1 309 1 FOFVF gag polyprotein - 369
297 36 37.1 309 2 S37028 exoM protein - Rhi 370
298 36 37.1 309 2 S39957 succinoglycan bios 371
299 36 37.1 309 2 C95977 glucosyltransferas 372
300 36 37.1 315 2 H97169 carbamoylphosphate 373
301 36 37.1 322 2 B81336 pseudouridylyate sy 374
302 36 37.1 323 2 G82135 probable lytic mur 375
303 36 37.1 333 2 C81245 ABC transporter, p 376
304 36 37.1 333 2 B82024 probable thiamin-b 377
305 36 37.1 337 2 F91191 probable LPS biosy 378
306 36 37.1 337 2 G86038 probable LPS biosy 379
307 36 37.1 350 2 B38535 A/G-specific adeni 380
308 36 37.1 350 2 H85953 adenine glycosylas 381
309 36 37.1 350 2 E91108 adenine glycosylas 382
310 36 37.1 372 2 T18697 hypothetical prote 383
311 36 37.1 373 2 F83020 UDP-glucose-heptos 384
312 36 37.1 376 1 F64705 conserved hypothet 385
313 36 37.1 376 2 F71815 hypothetical prote 386
314 36 37.1 382 2 G95896 probable transcrip 387
315 36 37.1 388 2 F89773 hypothetical prote 388
316 36 37.1 402 2 AD3370 NAD(FAD)-utilizing 389
317 36 37.1 409 2 E83179 hypothetical prote 390
318 36 37.1 419 2 E83904 hypothetical prote 391
319 36 37.1 429 2 S20050 transcription fact 392
320 36 37.1 445 2 H97075 4-aminobutyrate am 393
321 36 37.1 451 2 T49341 cytochrome P450 re 394

36 37.1 452 2 T05699 hypothetical prote 322
36 37.1 464 2 G70362 hypothetical prote 323
36 37.1 469 2 AD1926 hypothetical prote 324
36 37.1 471 2 B38637 Ras inhibitor (clo 325
36 37.1 489 2 AH0782 lysine-specific pe 326
36 37.1 491 2 H83658 hypothetical prote 327
36 37.1 505 2 H85361 leucyl aminopeptid 328
36 37.1 507 2 C82901 conserved hypothet 329
36 37.1 523 2 T18700 hypothetical prote 330
36 37.1 535 1 S76953 protein kinase (EC 331
36 37.1 587 2 S36231 probable beta-fruc 332
36 37.1 590 2 T02096 beta-fructofuranos 333
36 37.1 593 2 T01575 probable beta-fruc 334
36 37.1 648 1 WLWLC1 E1 protein - pygmy 335
36 37.1 664 2 H83962 serine/threonine p 336
36 37.1 666 2 T24170 hypothetical prote 337
36 37.1 701 1 FOFVLR gag polyprotein - 338
36 37.1 701 2 S11454 gag polyprotein - 339
36 37.1 701 2 F48613 gag polyprotein - 340
36 37.1 701 2 D48613 gag polyprotein - 341
36 37.1 701 2 S35430 gag polyprotein - 342
36 37.1 704 2 T01772 hypothetical prote 343
36 37.1 719 2 B95325 conserved hypothet 344
36 37.1 723 2 G84507 hypothetical prote 345
36 37.1 727 2 AC0076 beta-glucosidase (346
36 37.1 760 2 T34414 hypothetical prote 347
36 37.1 765 2 A69440 conserved hypothet 348
36 37.1 786 2 G88065 protein T16A1.2 [i 349
36 37.1 799 2 T02456 plasmin (EC 3.4.21 350
36 37.1 810 1 PLHU surface glycoprote 351
36 37.1 827 2 S29955 env polyprotein - 352
36 37.1 855 1 JQ2004 pyruvate dehydroge 353
36 37.1 882 2 G83018 glutamate receptor 354
36 37.1 896 2 A41273 glutamate receptor 355
36 37.1 906 2 A40222 glutamate receptor 356
36 37.1 906 2 S38723 glutamate receptor 357
36 37.1 906 2 S25852 glutamate receptor 358
36 37.1 908 2 T22376 hypothetical prote 359
36 37.1 980 2 T27342 hypothetical prote 360
36 37.1 1039 2 G83748 alpha-mannosidase 361
36 37.1 1068 1 A43322 1-phosphatidylinos 362
36 37.1 1068 1 I38110 1-phosphatidylinos 363
36 37.1 1079 1 TVFVMI gag-Rml- env polyp 364
36 37.1 1089 2 T31583 hypothetical prote 365
36 37.1 1124 2 B45557 HIV-1 retropepsin 366
36 37.1 1132 1 QSBPL host specificity p 367
36 37.1 1132 2 H90834 host specificity p 368
36 37.1 1137 2 B90734 probable host spec 369
36 37.1 1138 2 D85584 tat binding homolo 370
36 37.1 1190 2 T38636 virulence protein 371
36 37.1 1238 1 A40185 virulence sensor p 372
36 37.1 1238 1 S17944 virulence sensor p 373
36 37.1 1238 1 S17946 neurocan precursor 374
36 37.1 1257 2 S28764 microcystin synthe 375
36 37.1 1271 2 T43269 gag/pol polyprotei 376
36 37.1 1603 2 A48613 S-layer protein - 377
36 37.1 1616 2 T17884 probable polyketid 378
36 37.1 1733 2 D70887 polyketide synthas 379
36 37.1 1784 2 E86921 transcription init 380
36 37.1 2068 2 A47371 hypothetical prote 381
36 37.1 2962 2 T19756 Ran-binding protei 382
36 37.1 3224 1 S58884 elastin titin - hu 383
36 37.1 7962 2 I38346 Acl protein - mous 384
35.5 36.6 124 2 A54773 probable transposa 385
35.5 36.6 221 2 B60634 anthranilate dioxy 386
35.5 36.6 340 2 A83332 two-component sens 387
35.5 36.6 377 2 E84103 protein kinase - B 388
35.5 36.6 386 2 I39834 2,3-dihydroxybenzo 389
35.5 36.6 539 1 D69615 pyochelin biosynth 390
35.5 36.6 547 2 G83116 oligo-1,6-glucosid 391
35.5 36.6 558 2 S13579 hydrogenase (EC 1. 392
35.5 36.6 645 2 G72256 Down syndrome cell 393
1896 2 T08851 394

541	35	36.1	983	1	E45390	env polyprotein pr	614	34	35.1	129	2	H95402	hypothetical prote
542	35	36.1	992	2	A83324	probable sensor/re	615	34	35.1	145	2	B91252	hypothetical prote
543	35	36.1	1000	2	T30280	hypothetical prote	616	34	35.1	147	2	S57440	lydB protein - pha
544	35	36.1	1047	2	AI2002	hypothetical prote	617	34	35.1	151	2	S63748	HIV-1 retropepsin
545	35	36.1	1080	2	T00587	probable ubiquitin	618	34	35.1	151	2	AB0313	conserved hypothet
546	35	36.1	1086	1	B46335	HIV-1 retropepsin	619	34	35.1	153	2	A96751	hypothetical prote
547	35	36.1	1087	2	JQ1162	Pol protein - Maed	620	34	35.1	154	2	S28181	transcription fact
548	35	36.1	1094	2	F70697	probable arabinosy	621	34	35.1	158	2	AI0605	conserved hypothet
549	35	36.1	1101	1	B45390	HIV-1 retropepsin	622	34	35.1	158	2	E64823	hypothetical prote
550	35	36.1	1101	1	GNLJVS	HIV-1 retropepsin	623	34	35.1	158	2	E90745	probable sensory t
551	35	36.1	1118	1	SYBYCP	carbamoyl-phosphat	624	34	35.1	161	2	A69732	PBSX prophage ORF
552	35	36.1	1148	2	D82091	exodeoxyribonuclea	625	34	35.1	165	2	B87702	ribosomal protein
553	35	36.1	1148	2	T09073	splicing factor S1	626	34	35.1	184	2	H83409	hypothetical prote
554	35	36.1	1191	2	S76414	beta transducin-li	627	34	35.1	188	2	B72450	hypothetical prote
555	35	36.1	1259	4	GNHUL1	retrovirus-related	628	34	35.1	197	2	G82973	transcription regu
556	35	36.1	1275	2	I38588	reverse transcript	629	34	35.1	206	2	T08699	hypothetical prote
557	35	36.1	1275	2	S65824	reverse transcript	630	34	35.1	207	2	G89971	hypothetical prote
558	35	36.1	1275	2	B28096	line-1 protein ORF	631	34	35.1	209	1	XURTMC	conserved hypothet
559	35	36.1	1384	2	T26656	hypothetical prote	632	34	35.1	210	2	A69898	conjugative transf
560	35	36.1	1537	2	F86509	CT147 hypothetical	633	34	35.1	212	2	A39442	hypothetical prote
561	35	36.1	1537	2	C81558	conserved hypothet	634	34	35.1	214	2	A83416	hypothetical prote
562	35	36.1	1537	2	H72112	ct147 hypothetical	635	34	35.1	217	2	T15873	hypothetical prote
563	35	36.1	1657	2	T25421	hypothetical prote	636	34	35.1	219	2	C75637	3-oxoadipate CoA-t
564	35	36.1	1680	2	T01367	hypothetical prote	637	34	35.1	225	2	C83932	hypothetical prote
565	35	36.1	1930	2	F86200	protein F12K11.17	638	34	35.1	228	2	E89923	hypothetical prote
566	35	36.1	2207	2	T42759	Munc13-3 protein -	639	34	35.1	235	2	T19328	hypothetical prote
567	35	36.1	2869	2	T18518	apolipoprotein(a)	640	34	35.1	235	2	C83822	hypothetical prote
568	35	36.1	2948	2	T22664	hypothetical prote	641	34	35.1	235	2	G86567	lipote-protein li
569	35	36.1	4377	2	A55575	ankyrin 3, long sp	642	34	35.1	235	2	F72056	lipote-protein li
570	34.5	35.6	85	2	B91147	hypothetical prote	643	34	35.1	238	2	AH2178	orotidine 5' monop
571	34.5	35.6	85	2	F85992	hypothetical prote	644	34	35.1	240	2	AE1504	phosphoribosylform
572	34.5	35.6	85	2	C65120	hypothetical 10.0	645	34	35.1	241	2	T07067	beta-fructofuranos
573	34.5	35.6	119	2	PH1551	Ig H chain V regio	646	34	35.1	242	2	C83631	probable carbonic
574	34.5	35.6	171	2	T48286	hypothetical protei	647	34	35.1	245	2	T07071	beta-fructofuranos
575	34.5	35.6	211	2	T03636	GTP-binding protei	648	34	35.1	247	1	D24706	modulation protein
576	34.5	35.6	276	2	D64014	hypothetical prote	649	34	35.1	247	2	PQ0178	glyceraldehyde-3-p
577	34.5	35.6	291	2	I38167	hypothetical CpG-i	650	34	35.1	247	2	H95319	phosphoadenylyl-su
578	34.5	35.6	312	2	S34635	superantigen Mtv -	651	34	35.1	248	2	PQ0769	glycoprotein G - b
579	34.5	35.6	315	2	S23594	superantigen Mtv(M	652	34	35.1	250	2	PQ0768	glycoprotein G - b
580	34.5	35.6	315	2	JH0552	superantigen Mtv3	653	34	35.1	251	2	T11549	ATP binding protei
581	34.5	35.6	315	2	JH0554	superantigen Mtv13	654	34	35.1	252	2	T31439	probable cobyric a
582	34.5	35.6	315	2	S33143	superantigen Mtv -	655	34	35.1	255	1	RHRTT	thyloliberin precu
583	34.5	35.6	315	2	JH0551	superantigen Mtv1/	656	34	35.1	256	1	RTMST	thyloliberin precu
584	34.5	35.6	319	1	QQMVTM	superantigen Mtv(C	657	34	35.1	256	1	MFNZBR	matrix protein - b
585	34.5	35.6	320	1	QQEV3M	MTV superantigen -	658	34	35.1	257	1	MGNZBR	major surface glyc
586	34.5	35.6	320	2	S26174	superantigen Mtv17	659	34	35.1	261	2	AH2960	ABC transporter, m
587	34.5	35.6	320	2	S26387	superantigen Mtv9	660	34	35.1	261	2	F98322	nitrate transport
588	34.5	35.6	320	2	E26795	superantigen Mtv(B	661	34	35.1	264	2	AH2202	hypothetical prote
589	34.5	35.6	322	2	JH0706	superantigen Mtv7	662	34	35.1	265	2	E64128	lic-1 protein D -
590	34.5	35.6	322	2	S24574	superantigen Mtv(S	663	34	35.1	266	2	E71230	hypothetical prote
591	34.5	35.6	324	1	QQMV6M	MTV superantigen -	664	34	35.1	271	2	B83682	urease accessory p
592	34.5	35.6	325	2	S49019	superantigen - mou	665	34	35.1	272	2	T35231	hypothetical prote
593	34.5	35.6	325	2	S35303	superantigen Mtv -	666	34	35.1	274	2	E83444	probable transcrip
594	34.5	35.6	359	1	F64407	carotenoid biosynt	667	34	35.1	274	2	B75518	bacitracin resista
595	34.5	35.6	369	2	JC4292	probable transposa	668	34	35.1	280	2	T43011	suppressor protein
596	34.5	35.6	472	2	AG2417	hypothetical prote	669	34	35.1	281	2	G72680	hypothetical prote
597	34.5	35.6	482	2	H84539	probable homeodoma	671	34	35.1	283	2	T23785	hypothetical prote
598	34.5	35.6	524	2	T06021	hypothetical prote	672	34	35.1	286	2	AH2416	probable respirato
599	34.5	35.6	578	2	D82415	exoribonuclease II	673	34	35.1	286	2	T35001	SMR2 protein - Pod
600	34.5	35.6	719	2	S25237	homeotic protein H	674	34	35.1	287	2	S43852	glyceraldehyde-3-p
601	34	35.1	26	1	MEHBCF	melittin - little	675	34	35.1	288	2	S39888	transcription regu
602	34	35.1	56	2	AF2385	hypothetical prote	676	34	35.1	295	2	T07730	hypothetical prote
603	34	35.1	84	1	ZCBPG4	gene C protein - p	677	34	35.1	299	2	B83888	transcription regu
604	34	35.1	84	2	F84388	hypothetical prote	678	34	35.1	303	2	AI2520	hypothetical prote
605	34	35.1	90	2	AC2180	hypothetical prote	679	34	35.1	306	2	T25954	US10 protein - hum
606	34	35.1	92	2	AC2520	hypothetical prote	680	34	35.1	312	1	QQBE07	hypothetical prote
607	34	35.1	101	1	EEWT1	glutenin 1 - wheat	681	34	35.1	312	2	F86876	hypothetical prote
608	34	35.1	103	2	T07053	cysteine proteinas	682	34	35.1	313	2	AC0098	transcription acti
609	34	35.1	109	2	S69853	hypothetical prote	683	34	35.1	313	2	T49129	hypothetical prote
610	34	35.1	110	2	T47180	hypothetical prote	684	34	35.1	315	2	A40057	transforming growt
611	34	35.1	112	2	S75152	hypothetical prote	685	34	35.1	326	2	S73828	hypothetical prote
612	34	35.1	124	2	T46547	hypothetical prote	686	34	35.1	327	2	AC1615	glycine betaine/ca
613	34	35.1	125	2	F81197	holo- (acyl-carrier							

687	34	35.1	328	2	AE1252	glycine betaine/ca	760	34	35.1	452	2	D83609	probable glutamine
688	34	35.1	331	2	F82131	glyceraldehyde 3-p	761	34	35.1	454	2	T49300	hypothetical prote
689	34	35.1	333	1	DELOG3	glyceraldehyde-3-p	762	34	35.1	464	2	T17332	hypothetical prote
690	34	35.1	333	2	E97257	spore coat protein	763	34	35.1	468	2	T00794	hypothetical prote
691	34	35.1	337	1	DEJUGC	glyceraldehyde-3-p	764	34	35.1	470	2	AD0888	sufi protein [impo
692	34	35.1	337	1	DEPJG	glyceraldehyde-3-p	765	34	35.1	470	2	E91116	suppressor of ftsI
693	34	35.1	337	1	DESKG	glyceraldehyde-3-p	766	34	35.1	470	2	E85961	suppressor of ftsI
694	34	35.1	337	1	DEUSGM	glyceraldehyde-3-p	767	34	35.1	470	2	G65088	sufi protein precu
695	34	35.1	337	1	DEZMGC	glyceraldehyde-3-p	768	34	35.1	473	2	S04113	nitrogenase (EC 1.
696	34	35.1	337	2	S29814	glyceraldehyde-3-p	769	34	35.1	474	2	C35405	nitrogenase (EC 1.
697	34	35.1	337	2	S26863	glyceraldehyde-3-p	770	34	35.1	476	2	A83235	alkaline phosphata
698	34	35.1	338	1	A47183	hemoglobin precurs	771	34	35.1	477	2	H83588	probable MFS trans
699	34	35.1	338	2	T47218	glyceraldehyde-3-p	772	34	35.1	477	2	T25798	hypothetical prote
700	34	35.1	338	2	JN0452	glyceraldehyde-3-p	773	34	35.1	481	2	T15372	hypothetical prote
701	34	35.1	339	2	G64041	glyceraldehyde-3-p	774	34	35.1	485	2	S52411	ZNF165 protein - h
702	34	35.1	340	2	H83443	glycerol-3-phospha	775	34	35.1	489	1	C64984	lysine-specific pe
703	34	35.1	340	2	T51386	probable protein w	776	34	35.1	489	2	H91009	lysine-specific pe
704	34	35.1	344	2	A95402	probable desaturas	777	34	35.1	489	2	B85854	lysine-specific pe
705	34	35.1	347	2	C90598	hypothetical prote	778	34	35.1	491	2	H84477	probable PttA-like
706	34	35.1	351	1	A48763	transcription fact	779	34	35.1	492	2	D71510	probable oligopept
707	34	35.1	352	2	G84077	hypothetical prote	780	34	35.1	492	2	G81668	peptide ABC transp
708	34	35.1	355	2	T20825	hypothetical prote	781	34	35.1	499	2	C75251	occludin - chicken
709	34	35.1	357	2	T25777	hypothetical prote	782	34	35.1	504	2	A49467	phosphopyruvate hy
710	34	35.1	359	1	LQBP37	DNA ligase (ATP) (783	34	35.1	513	2	AE1921	hypothetical prote
711	34	35.1	359	2	S52650	omega-3 fatty acid	784	34	35.1	515	2	G91271	hypothetical prote
712	34	35.1	360	2	S38570	glyceraldehyde-3-p	785	34	35.1	515	2	G86112	hypothetical prote
713	34	35.1	361	1	F65012	hypothetical prote	786	34	35.1	515	2	S56392	hypothetical 54.7K
714	34	35.1	361	2	A91037	probable peptidase	787	34	35.1	523	2	T30091	hypothetical prote
715	34	35.1	361	2	C85881	probable peptidase	788	34	35.1	524	2	E71881	hypothetical prote
716	34	35.1	362	2	T49528	hypothetical prote	789	34	35.1	525	2	D90048	hypothetical prote
717	34	35.1	365	2	B83829	glycine oxidase (s	790	34	35.1	531	1	SYNCYT	tyrosine-trNA liga
718	34	35.1	366	2	AH2429	hypothetical prote	791	34	35.1	534	2	T41081	hypothetical prote
719	34	35.1	367	1	I69653	UDPgallactopyranose	792	34	35.1	536	2	AB1983	hypothetical prote
720	34	35.1	368	2	G81289	UDPgallactopyranose	793	34	35.1	544	2	T38469	conserved hypothet
721	34	35.1	377	2	S30088	aciA protein - Eme	794	34	35.1	549	1	S19095	transcription fact
722	34	35.1	379	2	I48133	ubiquinol-cytochro	795	34	35.1	553	2	AH1830	serine/threonine k
723	34	35.1	379	2	I48132	ubiquinol-cytochro	796	34	35.1	557	2	AF3091	hypothetical prote
724	34	35.1	379	2	I48134	ubiquinol-cytochro	797	34	35.1	557	2	D98195	hypothetical prote
725	34	35.1	379	2	I48134	ubiquinol-cytochro	798	34	35.1	559	2	AI0135	DNA repair protein
726	34	35.1	382	2	T34789	hypothetical prote	799	34	35.1	562	2	S55516	glucan 1,3-beta-gl
727	34	35.1	384	2	T51423	formate dehydrogen	800	34	35.1	564	2	S41974	DNA ligase (ATP) (
728	34	35.1	384	2	S34816	nitrogenase cofact	801	34	35.1	566	2	S50979	RHC21 protein - ye
729	34	35.1	386	2	F95911	probable hydrocarb	802	34	35.1	568	2	T05218	hypothetical prote
730	34	35.1	389	2	B69096	corrinoid/iron-sul	803	34	35.1	571	2	T333941	hypothetical prote
731	34	35.1	389	2	S73910	probable serine/th	804	34	35.1	573	2	A82231	transport ATP-bind
732	34	35.1	390	2	A27512	transforming growt	805	34	35.1	580	2	S49308	beta-fructofuranos
733	34	35.1	390	2	I46463	transforming growt	806	34	35.1	580	2	D69645	myo-inositol catab
734	34	35.1	391	2	S01413	transforming growt	807	34	35.1	580	2	H69793	adenine deaminase
735	34	35.1	391	2	E83151	hypothetical prote	808	34	35.1	586	1	JVBBP3	DNA maturase B - p
736	34	35.1	391	2	D83777	hypothetical prote	809	34	35.1	586	1	JVBBP7	DNA maturase B - p
737	34	35.1	399	2	F72417	hypothetical prote	810	34	35.1	592	2	S61503	beta-fructofuranos
738	34	35.1	400	2	S76929	hypothetical prote	811	34	35.1	592	2	S56681	beta-fructofuranos
739	34	35.1	400	2	AF0291	probable drug resi	812	34	35.1	592	2	JQ0991	beta-fructofuranos
740	34	35.1	400	2	T18570	probable malate de	813	34	35.1	603	2	T40528	palmitoyl-protein
741	34	35.1	408	2	T50876	hypothetical membr	814	34	35.1	605	2	T33750	hypothetical prote
742	34	35.1	411	2	AG0045	probable iron-sulf	815	34	35.1	608	2	C96919	glutamine-fructose
743	34	35.1	412	2	D95941	conserved hypothet	816	34	35.1	608	2	T02684	MYB-related transc
744	34	35.1	417	2	AF3448	hypothetical membr	817	34	35.1	618	2	AI0171	probable exported
745	34	35.1	418	2	S23675	contrapsin-related	818	34	35.1	620	2	I57937	dopamine transport
746	34	35.1	418	2	S02138	arginine deiminase	819	34	35.1	620	2	A48980	dopamine transport
747	34	35.1	418	2	S76699	hypothetical prote	820	34	35.1	625	1	W1WLR1	E1 protein - rhesu
748	34	35.1	427	2	AE1305	conserved hypothet	821	34	35.1	633	2	B70946	NADH2 dehydrogenas
749	34	35.1	427	2	AE1677	conserved hypothet	822	34	35.1	639	2	T50148	probable ser/thr p
750	34	35.1	430	2	S05481	keratin 18, type I	823	34	35.1	647	1	I58353	LIM protein kinase
751	34	35.1	431	2	AE0551	phosphate regulon	824	34	35.1	647	1	JP0078	LIM protein kinase
752	34	35.1	432	2	A83060	hypothetical prote	825	34	35.1	662	2	T41442	omnipotent nonsens
753	34	35.1	435	2	T06210	probable glutamate	826	34	35.1	662	2	T51948	omnipotent nonsens
754	34	35.1	438	2	G87337	membrane protein,	827	34	35.1	662	2	JC7906	sucrose 1F-fructos
755	34	35.1	440	2	A83435	ATP synthase in ty	828	34	35.1	664	1	WMBBEH	infected cell prot
756	34	35.1	443	2	T05132	hypothetical prote	829	34	35.1	668	2	B96999	methyl-accepting c
757	34	35.1	445	2	B75519	probable ubiquinol	830	34	35.1	684	2	F70810	hypothetical prote
758	34	35.1	451	2	B90312	hypothetical prote	831	34	35.1	687	1	B69649	beta-galactosidase
759	34	35.1	451	2	E72398	hypothetical prote	832	34	35.1	687	2	B70515	hypothetical prote

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 44.3%; Score 43; DB 2; Length 217;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSRE 15
||| ||| | : |
Db 87 GPRLRQWPLSKE 98

RESULT 10
A83862
Initiation of chromosome replication dnaD [imported] - Bacillus halodurans (strain C-125
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A83862
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83862
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <STO>
A;Cross-references: UNIPROT:Q9KC77; UNIPARC:UPI00000C3C7C; GB:AP001512; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: dnaD

Query Match 44.3%; Score 43; DB 2; Length 233;
Best Local Similarity 43.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSREHT 17
| : | | | | : : |
Db 144 IEGETLSMWIDQDQHT 159

RESULT 11
S07740
hypothetical protein 8 - Paramecium tetraurelia mitochondrion
C;Species: mitochondrion Paramecium tetraurelia
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: S07740
R;Pritchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummings,
Nucleic Acids Res. 18, 173-180, 1990
A;Title: Nucleotide sequence of the mitochondrial genome of Paramecium.
A;Reference number: S07725; MUID:90174913; PMID:2308823
A;Accession: S07740
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-241 <PRI>
A;Cross-references: UNIPROT:P15609; UNIPARC:UPI000013B890; EMBL:X15917; NID:g13256; PID:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC6
A;Start codon: ATT
C;Superfamily: mitochondrial ribosomal protein S18, paramecium type
C;Keywords: mitochondrion

Query Match 44.3%; Score 43; DB 2; Length 241;
Best Local Similarity 41.2%; Pred. No. 25;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSREHTS 18
: | | | | : : | | : :
Db 26 VKGPTIEKFLKFEYNA 42

RESULT 12
S73999
hypothetical protein yaaC homolog VxpsPT7_orf269 - Mycoplasma pneumoniae (strain ATCC 29
N;Alternate names: hypothetical protein VxpsPT7_orf269
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S73999
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73999
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-269 <HIM>
A;Cross-references: UNIPROT:P75587; UNIPARC:UPI00001338E1; EMBL:AE000062; GB:U00089; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: uncharacterized conserved protein HI0963

Query Match 44.3%; Score 43; DB 2; Length 269;
Best Local Similarity 58.3%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQWLK 12
| : | : | | | |
Db 124 TLSSSTIRQWLK 135

RESULT 13
S30484
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30484
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30484
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
C;Superfamily: pol polyprotein

Query Match 44.3%; Score 43; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSRE 15
| | | | | | : |
Db 31 GPKLRQWPLSKE 42

RESULT 14
S30483
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30483
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30483
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111

C;Superfamily: pol polypeptide

Query Match 44.3%; Score 43; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSR 15
||| ||| ||| |||
DB 31 GPKLRQWPLSK 42

RESULT 15
S59644
sister chromatid cohesion molecule Mis4p - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 14-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C;Accession: T38603; T43392; S59644
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, July 1995

A;Reference number: Z21731
A;Accession: T38603
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1583 <DE2>

A;Cross-references: UNIPROT:Q09725; UNIPARC:UPI000012F161; EMBL:Z50113; NID:g914878; PID
A;Experimental source: strain 972h-; cosmid c31A2
R;Furuya, K.; Takahashi, K.; Yanagida, M.
submitted to the EMBL Data Library, August 1998
A;Description: Faithful anaphase is ensured by Mis4, a sister chromatid cohesion molecule

A;Reference number: Z22478
A;Accession: T43392
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 1-1583 <FUR>

A;Cross-references: UNIPARC:UPI000012F161; EMBL:AB016866; PIDN:CAB19489.1
C;Genetics:
A;Gene: mis4; SPAC31A2.05c
A;Map position: 1
A;Introns: 33/1; 98/2; 543/3; 699/3; 1294/2; 1339/3; 1558/3

Query Match 44.3%; Score 43; DB 2; Length 1583;
Best Local Similarity 46.2%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSR 16
||| ||| ||| |||
DB 1483 GPTITGWMKKLDH 1495

RESULT 16
AC0189
probable exported protein YPO1551 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0189
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0189
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <KUR>
A;Cross-references: UNIPROT:Q8ZFX2; UNIPARC:UPI00000DCBC5; GB:AL590842; PIDN:CAC90374.1;
C;Genetics:
A;Gene: YPO1551

Query Match 43.8%; Score 42.5; DB 2; Length 275;
Best Local Similarity 47.6%; Pred. No. 34;
Matches 10; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 TIKGPT---LRQWLKSR 18
|: || | ||| |||
DB 133 TVAGKTLMALEQWLHQLPHTS 153

RESULT 17

E71171
hypothetical protein PH0569 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: E71171

R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki,
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kuchida, N.; Oguchi,
DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: E71171

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-114 <KAW>

A;Cross-references: UNIPROT:O58304; UNIPARC:UPI0000062E54; GB:AP000002; NID:g3236129; P

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:
A;Gene: PH0569

Query Match 43.3%; Score 42; DB 2; Length 114;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLK 12
:|| ||| |||
DB 47 VKGDTLKVWLK 57

RESULT 18

T20567

hypothetical protein F08A10.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T20567

R;Kershaw, J.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z19293

A;Accession: T20567

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-171 <WIL>

A;Cross-references: UNIPROT:Q19186; UNIPARC:UPI000017B8BA; EMBL:Z75534; PIDN:CAA99825.1

A;Experimental source: clone F08A10

C;Genetics:
A;Gene: CESP:F08A10.1

A;Map position: 1

A;Introns: 26/3; 64/1; 91/3; 118/3

Query Match 43.3%; Score 42; DB 2; Length 171;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSR 17
||| ||| ||| |||
DB 103 GPSLRPFLNSGNHT 116

RESULT 19

D72456

probable glutamyl-tRNA reductase APE2296 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: D72456

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain 122
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: D72456
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-416 <KAW>
A;Cross-references: UNIPROT:Q9Y9J2; UNIPARC:UPI000005E2B8; DDBJ:AP0000064; NID:g5105945;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2296
C;Superfamily: glutamyl-tRNA reductase

Query Match 43.3%; Score 42; DB 2; Length 416;
Best Local Similarity 57.1%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRE 15
: | | | | | | |
Db 105 VLQVRRRAWLKSRE 118

RESULT 20

S71157
cytochrome c biogenesis protein 454 - evening primrose mitochondrion
C;Species: mitochondrion Oenothera berteriana (evening primrose)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
A;Accession: S71157; S55283; S42984
R;Gruska, I.; Jakabsons, W.; Schuster, W.
Mol. Gen. Genet. 247, 529-536, 1995
A;Title: Oenothera mitochondrial orf454, a gene involved in cytochrome c biogenesis cor
A;Reference number: S55283; MUID:95327048; PMID:7603431
A;Accession: S71157
A;Molecule type: mRNA
A;Residues: 1-438 <GRU>
A;Cross-references: UNIPROT:Q35213; UNIPARC:UPI000017CA75; EMBL:X78036
A;Note: differences are due to RNA editing; premature stop codon is due to RNA editing
A;Accession: S55283
A;Molecule type: DNA
A;Residues: 1-16, 'PR', 19-34, 'P', 36-39, 'SS', 42-48, 'P', 50, 'PS', 53, 'P', 55-108, 'L', 110-130, '
A;Cross-references: UNIPARC:UPI00000964AA; EMBL:X78036; NID:g459536; PIDN:CAA54966.1; PI
C;Genetics:
A;Genome: mitochondrion
A;Introns: 256/2
C;Keywords: mitochondrion; RNA editing

Query Match 43.3%; Score 42; DB 2; Length 438;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 LRQWLKSREHTS 18
| : | : | : | : |
Db 342 LHRWVKNREHN 353

RESULT 21

C86160
hypothetical protein F22D16.3 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86160
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86160
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-519 <STO>
A;Cross-references: UNIPROT:Q9SRY9; UNIPARC:UPI00000A8372; GB:AE005172; NID:g6056405; P;
C;Genetics:
A;Map position: 1

Query Match 43.3%; Score 42; DB 2; Length 519;
Best Local Similarity 40.0%; Pred. No. 82;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREH 16
| : | : | : | : | : |
Db 498 IRGHTMVEWLKAKQN 512

RESULT 22

T12095
beta-fructofuranosidase (EC 3.2.1.26), cell wall - fava bean
N;Alternate names: cell wall invertase II
C;Species: Vicia faba (fava bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T12095
R;Weber, H.; Borisjuk, L.; Heim, U.; Buchner, P.; Wobus, U.
Plant Cell 7, 1835-1846, 1995
A;Title: Seed coat-associated invertases of Fava bean control both unloading and storage
A;Reference number: Z17416; MUID:96093423; PMID:8535137
A;Accession: T12095
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-581 <WEB>
A;Cross-references: UNIPROT:Q43856; UNIPARC:UPI00000A5748; EMBL:Z35163; NID:g861156; PI
A;Experimental source: strain minor; cultivar Fribo; seed coat; clone VfCWINV2
C;Genetics:
A;Gene: CWINV2
C;Function:
A;Description: hydrolyzes terminal non-reducing beta-D-fructofuranoside residues in bet
C;Superfamily: beta-fructofuranosidase
C;Keywords: cell wall; glycoprotein; glycosidase; hydrolase

Query Match 43.3%; Score 42; DB 2; Length 581;
Best Local Similarity 58.3%; Pred. No. 93;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PTLRQWLKSREH 16
| | | : | : | : | : |
Db 160 PFLREWIKSPEN 171

RESULT 23

S75447
proline-tRNA ligase - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sll1425
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75447
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75447
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-617 <KAN>
A;Cross-references: UNIPROT:P73942; UNIPARC:UPI0000136614; EMBL:D90911; GB:AB001339; NI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: proS
A;Start codon: GTG
C;Superfamily: proline-tRNA ligase

Query Match 43.3%; Score 42; DB 2; Length 617;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLRQWLKSREHTS 18
||| |:|: ||
Db 604 PTLTAWIKAEKTS 617

RESULT 24
GNLJG4
HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (African green monkey i
N;Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: simian immunodeficiency virus, SIV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
C;Accession: B30045
R;Fukasawa, M.; Miura, T.; Hasegawa, A.; Morikawa, S.; Tsujimoto, H.; Miki, K.; Kitamura
Nature 333, 457-461, 1988
A;Title: Sequence of simian immunodeficiency virus from African green monkey, a new memb
A;Reference number: A30045; MUID:88232906; PMID:3374586
A;Accession: B30045
A;Molecule type: DNA
A;Residues: 1-1061 <FUK>
A;Cross-references: UNIPARC:UPI0000131F18; EMBL:X07805; NID:g61748; PID:g1335593
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse
F;111-210/Product: retropepsin #status predicted <RTP>
F;134/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 43.3%; Score 42; DB 1; Length 1061;
Best Local Similarity 53.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRE 15
:|:|:|:|:|
Db 229 RGPVCRQWPLSKE 241

RESULT 25
H70819
probable polyketide synthase - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70819
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70819
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2108 <COL>
A;Cross-references: UNIPROT:O53901; UNIPARC:UPI00000D3B08; GB:AL022000; GB:AL123456; NID
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: pks5
C;Superfamily: mycocerosic acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I ho
nase homology; [acyl-carrier-protein] S-malonyltransferase homology
C;Keywords: carrier protein
F;36-434/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F;546-826/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F;1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>
F;1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>
F;2029-2094/Domain: acyl carrier protein homology <ACPI>

Query Match 43.3%; Score 42; DB 2; Length 2108;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREH 16
:|:|:|:|:|
Db 1047 VDGAEVROWLDKRGH 1061

RESULT 26
T03098
p97 protein - Toxoplasma gondii
C;Species: Toxoplasma gondii
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03098
R;Matsuura, T.; Kasper, L.
Mol. Biochem. Parasitol. 90, 403-413, 1997
A;Title: Molecular analysis and characterization of a protein involved in the replicati
A;Reference number: Z14838; MUID:98135655; PMID:9476788
A;Accession: T03098
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-877 <MAT>
A;Cross-references: UNIPROT:O15644; UNIPARC:UPI000007FE49; EMBL:AF005059; NID:g2581824;
A;Experimental source: strain RH
C;Function:
A;Description: involved in replication of intracellular Toxoplasma gondii
C;Superfamily: Toxoplasma gondii p97 protein

Query Match 42.8%; Score 41.5; DB 2; Length 877;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 4 GP---TLRQWLKSREH 16
||| |:|:|:|
Db 72 GPVPNTFQWLQQNEH 87

RESULT 27
T06920
ribosomal protein L28 - Cyanophora paradoxa cyanelle
C;Species: cyanelle Cyanophora paradoxa
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06920
R;Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A;Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A;Reference number: Z15840
A;Accession: T06920
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-70 <STI>
A;Cross-references: UNIPROT:P48129; UNIPARC:UPI0000133A7D; EMBL:U30821; NID:g1016083; P
A;Experimental source: strain Pringsheim LB555
C;Genetics:
A;Gene: rpl28
A;Genome: cyanelle
C;Superfamily: Escherichia coli ribosomal protein L28
C;Keywords: cyanelle; ribosome

Query Match 42.3%; Score 41; DB 2; Length 70;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
|:|:|:|:|:|
Db 38 IWSPTLKRWLKLOVST 53

RESULT 28
A97524
hypothetical protein AGR_C_2500 [imported] - Agrobacterium tumefaciens (strain C58, Ce
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97524

111

R.; Klenk, H. P.; Clayton, R. A.; Tomb, J. F.; White, O.; Nelson, K. E.; Ketchum, K. A.; Dodson, J.; Fleischmann, R. D.; Quackenbush, J.; Lee, N. H.; Sutton, G. G.; Gill, S.; Kirkness, E. F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69449
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-397 <KLE>
A;Cross-references: UNIPROT:O28672; UNIPARC:UPI0000056C7D; GB:AE000782; NID
C;Genetics:
A;Gene: trpB-2
C;Function:
A;Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan
A;Pathway: tryptophan biosynthesis
A;Note: cofactor pyridoxal phosphate; last step in pathway
C;Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology
C;Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate; trypt
F;12-393/Domain: tryptophan synthase beta chain homology <TRPB>
F;93/Active site: His #status predicted
F;94/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 42.3%; Score 41; DB 2; Length 397;
Best Local Similarity 63.6%; Pred. No. 88;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 LRQWLKSREHT 17
|||:::|
Db 181 LRDWVESFEHT 191

RESULT 38
S29127
carboxypeptidase A (EC 3.4.17.1) CPA1 precursor - human
N;Alternate names: pancreatic carboxypeptidase A1
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1994 #sequence revision 19-Jan-1996 #text change 09-Jul-2004
C;Accession: S29127; A34205; S08253; S02810; S71394; S02811
R;Catasus, L.; Villegas, V.; Pascual, R.; Aviles, F.X.; Wicker-Planquart, C.; Puigserver
Biochem. J. 287, 299-303, 1992
A;Title: cDNA cloning and sequence analysis of human pancreatic procarboxypeptidase A1.
A;Reference number: S29127; MUID:93038569; PMID:1417781
A;Accession: S29127
A;Molecule type: mRNA
A;Residues: 1-419 <CAT>
A;Cross-references: UNIPROT:P15085; UNIPARC:UPI000004D332; EMBL:X67318; NID:g35329; PIDN
R;Stewart, E.A.; Craik, C.S.; Hake, L.; Bowcock, A.M.
Am. J. Hum. Genet. 46, 795-800, 1990
A;Title: Human carboxypeptidase A identifies a BgIII RFLP and maps to 7q31-qter.
A;Reference number: A34205; MUID:90196012; PMID:1969228
A;Accession: A34205
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 330-396 <STE>
A;Cross-references: UNIPARC:UPI0000172A5F
A;Note: the authors translated the codon CTG for residue 391 as Val
R;Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.
FEBS Lett. 261, 179-183, 1990
A;Title: Further studies on the human pancreatic binary complexes involving procarboxype
A;Reference number: S08253; MUID:90169111; PMID:2307232
A;Accession: S08253
A;Molecule type: protein
A;Residues: 17-43; 'XXX', 114-135 <MOU>
A;Cross-references: UNIPARC:UPI0000172A60; UNIPARC:UPI00000172A61
R;Pascual, R.; Burgos, F.J.; Salva, M.; Soriano, F.; Mendez, E.; Aviles, F.X.
Eur. J. Biochem. 179, 609-616, 1989
A;Title: Purification and properties of five different forms of human procarboxypeptidas
A;Reference number: S02809; MUID:89153096; PMID:2920728
A;Accession: S02810
A;Molecule type: protein
A;Residues: 17-42 <PAS>
A;Cross-references: UNIPARC:UPI00000172A62

R;Laethem, R.M.; Blumenkopf, T.A.; Cory, M.; Elwell, L.; Moxham, C.P.; Ray, P.H.; Walton
Arch. Biochem. Biophys. 332, 8-18, 1996
A;Title: Expression and characterization of human pancreatic preprocarboxypeptidase A1
A;Reference number: S71394; MUID:96400327; PMID:8806703
A;Accession: S71394
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-419 <LAE>
A;Cross-references: UNIPARC:UPI000004D332
C;Genetics:
A;Gene: GDB:CPA1; CPA
A;Cross-references: GDB:120597; OMIM:114850
A;Map position: 7q32-7qter
C;Superfamily: carboxypeptidase
C;Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein digestion; zinc
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-110/Domain: activation peptide #status predicted <ACP>
F;11-419/Product: carboxypeptidase A isozyme 1 #status predicted <MAT>
F;179,182,306/Binding site: zinc (His, Glu, His) #status predicted
F;248-271/Disulfide bonds: #status predicted
F;358,380/Active site: Tyr, Glu #status predicted

Query Match 42.3%; Score 41; DB 1; Length 419;
Best Local Similarity 53.8%; Pred. No. 94;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 TLRQWLKSREHTS 18
|||:::|
Db 232 TNRMWRKTRSHTA 244

RESULT 39
F70876
probable papA3 protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70876
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqaures, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70876
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-472 <COL>
A;Cross-references: UNIPROT:O50438; UNIPARC:UPI00000D1152; GB:AL010186; GB:AL123456; NII
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: papA3

Query Match 42.3%; Score 41; DB 2; Length 472;
Best Local Similarity 38.9%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSREHTS 18
|:::|:::|
Db 224 TVESQVRAWTKFAEGTN 241

RESULT 40
S35122
site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) LlaI - Lactococcus
N;Alternate names: type II modification methylase LlaI
C;Species: Lactococcus lactis
C;Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 31-Dec-2004
C;Accession: S35122; S77702; A47029
R;Hill, C.; Miller, L.A.; Klaenhammer, T.R.
J. Bacteriol. 173, 4363-4370, 1991
A;Title: In vivo genetic exchange of a functional domain from a type II A methylase bet
A;Reference number: A47029; MUID:91294179; PMID:1906061

A;Accession: S35122
A;Molecule type: DNA
A;Residues: 1-622 <HIL>
A;Cross-references: UNIPROT:P35516; UNIPARC:UPI000017AC4E; EMBL:M77136
A;Experimental source: bacteriophage resistance plasmid pTR2030
A;Note: the sequence of residues 469 and 470 is interchanged in the authors' translation
A;Note: sequence extracted from NCBI backbone (NCBIN:41635, NCBIP:41636)
R;Klaenhammer, T.R.
submitted to the EMBL Data Library, November 1994
A;Reference number: S77702
A;Accession: S77702
A;Molecule type: DNA
A;Residues: 1-248,'G',250-622 <KIA>
A;Cross-references: UNIPARC:UPI000012F7DD; EMBL:U17233; NID:g639886; PIDN:AAA65073.1; PI
C;Genetics:
A;Gene: L1aI
A;Genome: plasmid
A;Superfamily: modification methylase (adenine-specific), M.StsI type
C;Keywords: methyltransferase; restriction modification system; S-adenosylmethionine

Query Match 42.3%; Score 41; DB 2; Length 622;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSREHT 17
| | | | | | | |
Db 81 GKTPFQWLNEREYT 94

RESULT 41
S46347
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SAB-1)
C;Species: simian immunodeficiency virus SIVagm
A;Variety: isolate SAB-1
C;Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C;Accession: S46347
R;Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;
EMBO J. 13, 2935-2947, 1994
A;Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A;Reference number: S46335; MUID:94298785; PMID:8026477
A;Accession: S46347
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1039 <JIN>
A;Cross-references: UNIPARC:UPI00001096DD; EMBL:U04005; NID:g466229; PIDN:AAA21505.1; PI
A;Experimental source: isolate SAB-1; sabaesus monkey
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
A;Note: this reading frame extends between two stop codons and does not begin with a sta
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein

Query Match 42.3%; Score 41; DB 2; Length 1039;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRE 15
| | | | | | | |
Db 205 GPRIRQWLPSKE 216

RESULT 42
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; UNIPARC:UPI00000A945F; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190

Query Match 42.3%; Score 41; DB 2; Length 1123;
Best Local Similarity 58.3%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IXGPTLRQWLKS 13
: : | | | | | |
Db 200 VQPPYKRWLSS 211

RESULT 43
S73786
hypothetical protein A19_orf1140 - Mycoplasma pneumoniae (strain ATCC 29342)
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S73786
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73786
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1140 <HIM>
A;Cross-references: UNIPROT:P75405; UNIPARC:UPI000013A8AC; EMBL:AE000045; GB:U00089; NI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3

Query Match 42.3%; Score 41; DB 2; Length 1140;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSREH 16
| | | | | | | |
Db 1116 GWLHRWRKSRKH 1128

RESULT 44
T31091
hypothetical protein wbbK [imported] - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31091
R;Saigi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A.;
J. Bacteriol. 181, 1883-1891, 1999
A;Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.
A;Reference number: Z20974; MUID:99173913; PMID:10074083
A;Accession: T31091
A;Status: preliminary; translated from GB/EMBL/DDAJ
A;Molecule type: DNA
A;Residues: 1-1191 <SAI>
A;Cross-references: UNIPROT:O52484; UNIPARC:UPI00000BA53A; EMBL:AF038816; NID:g2828669;
C;Genetics:
A;Gene: wbbK

Query Match 42.3%; Score 41; DB 2; Length 1191;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 PTLRQWLKSREHTS 18
| | | | | | | |
Db 491 PELTQWLREARETA 504

RESULT 45
T13423
hypothetical protein 30B8.4 - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13423
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
A:Reference number: Z17668
A:Accession: T13423
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-3345 <MUR>
A:Cross-references: UNIPROT:O46074; UNIPARC:UPI00000796D3; EMBL:AL009195; NID:e1355203;
C:Genetics:
A:Cross-references: FlyBase:FBgn0000377
A:Map position: X
A:Introns: 51/3; 159/1; 476/1; 526/1; 1465/1; 1826/3; 1947/3; 2081/1; 2196/3; 3007/3
A:Note: EG:30B8.4

Query Match 42.3%; Score 41; DB 2; Length 3345;
Best Local Similarity 58.3%; Pred. No. 9.1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKX 13
:| | | | |
Db 2203 VKNPKLEQWLAS 2214

RESULT 46
D72305
hypothetical protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72305
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72305
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: UNIPROT:Q9X0A5; UNIPARC:UPI00000C1322; GB:AE001763; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1012

Query Match 41.2%; Score 40; DB 2; Length 158;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 KGPTLRQWLKSR 15
| | | | |
Db 145 KAETLRKWLNERK 157

RESULT 47
T49611
hypothetical protein B3E4.240 [imported] - *Neurospora crassa*
C:Species: *Neurospora crassa*
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49611
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49611
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-185 <SCH>
A:Cross-references: UNIPROT:Q9PSV8; UNIPARC:UPI000017B4A4; EMBL:AL3555931; GSPDB:GN00116,
A:Experimental source: BAC clone B3E4; strain OR74A
C:Genetics:
A:Gene: NCSP:B3E4.240
A:Map position: 6

Query Match 41.2%; Score 40; DB 2; Length 185;
Best Local Similarity 33.3%; Pred. No. 55;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQWLKSR 15
|:|:|:|:|:|:
Db 84 TLQGPSAQDWARHD 98

RESULT 48
T23485
hypothetical protein K08F4.11 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T23485
R:Hembry, C.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19746
A:Accession: T23485
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-200 <WIL>
A:Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN000022;
A:Experimental source: clone K08F4
C:Genetics:
A:Gene: CESP:K08F4.11
A:Map position: 4
A:Introns: 45/1; 76/1; 111/3
C:Superfamily: glutathione transferase

Query Match 41.2%; Score 40; DB 2; Length 200;
Best Local Similarity 43.8%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSR 17
|:|:|:|:|:|:
Db 183 IETPKLKEWLAKRPET 198

RESULT 49
T37464
probable glutathione transferase (EC 2.5.1.18) GST3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37464
R:Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.
submitted to the EMBL Data Library, June 1997
A:Description: Paraquat mediates differential gene expression in *C. elegans*.
A:Reference number: Z21702
A:Accession: T37464
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-207 <TAW>
A:Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AAB65419
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: GST3
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 41.2%; Score 40; DB 2; Length 207;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSR 17
|:|:|:|:|:|:
|

THIS PAGE BLANK (UPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 127.385 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-80
Perfect score: 97
Sequence: 1 TIKGPTLRQWLKSRHTS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		Match	Length	DB	ID	Description
	Score	%					
1	49	50.5	122	2	Q8ZC84	YERPE	Q8zc84 yersinia pe
2	49	50.5	122	2	Q66DR4	YERPS	Q66dr4 yersinia ps
3	48.5	50.0	718	2	Q5AQW2	EMENI	Q5aqw2 aspergillus
4	48	49.5	148	2	Q4UK00	RICFE	Q4uk00 rickettsia
5	48	49.5	326	2	P95613	RHIGA	P95613 rhizobium g
6	48	49.5	992	2	Q4P0J6	USTMA	Q4p0j6 ustilago ma
7	47	48.5	327	1	Y745	HELPI	Q7z198 helicobacte
8	47	48.5	727	2	Q750V6	ASHGO	Q750v6 ashbya goss
9	47	48.5	1794	2	Q55ZA2	CRYNE	Q55za2 cryptococcu
10	47	48.5	1811	2	Q5KNL7	CRYNE	Q5knl7 cryptococcu
11	46.5	47.9	332	1	GPDA2	MYCPA	P61744 mycobacteri
12	46	47.4	126	2	Q6YUC3	ORYSA	Q6yuc3 oryza sativ
13	46	47.4	365	2	Q6CL14	YARLI	Q6cli4 yarrowia li
14	46	47.4	402	2	Q82R87	STRAW	Q82r87 streptomyce
15	46	47.4	815	2	Q4KS46	9VIRU	Q4ks46 orange-spot
16	46	47.4	941	2	Q8QUJ6	9VIRU	Q8quj6 infectious
17	45	46.4	186	2	Q980N7	SULSO	Q980n7 sulfolobus
18	45	46.4	243	2	Q8XR40	RALSO	Q8xr40 ralstonia s
19	45	46.4	286	2	Q8T462	DROME	Q8t462 drosophila
20	45	46.4	302	2	Q742B3	MYCPA	Q742b3 mycobacteri
21	45	46.4	351	2	Q7VI26	HELHP	Q7vi26 helicobacte
22	45	46.4	789	2	Q6N3E8	REHOP	Q6n3e8 rhodopseudo
23	45	46.4	1126	2	Q4QH74	LEIMA	Q4qh74 leishmania
24	45	46.4	1361	2	Q5TV94	ANOGA	Q5tv94 anopheles g
25	45	46.4	1431	2	Q7PS78	ANOFA	Q7ps78 anopheles g
26	44.5	45.9	642	2	Q4KFC8	PSEF5	Q4kfc8 pseudomonas
27	44	45.4	131	2	Q4JW69	CORJK	Q4jw69 corynebacte
28	44	45.4	161	2	P93490	PEA	P93490 pisum sativ
29	44	45.4	298	2	Q5YPT6	NOCFA	Q5ypt6 nocardia fa
30	44	45.4	302	2	Q4K6B2	PSEF5	Q4k6b2 pseudomonas
31	44	45.4	313	2	P90433	SIVCZ	P90433 chimpanzee

32	44	45.4	332	2	Q4NFI4	9MICC	Q4nfi4 arthrobacte
33	44	45.4	347	2	Q8G7M6	BIFLO	Q8g7m6 bifidobacte
34	44	45.4	371	2	Q53MP5	ORYSA	Q53mp5 oryza sativ
35	44	45.4	377	2	Q82PX5	STRAW	Q82px5 streptomyce
36	44	45.4	472	2	Q4PED7	USTMA	Q4ped7 ustilago ma
37	44	45.4	500	2	Q6AKH3	DESBS	Q6akh3 desulfotale
38	44	45.4	648	2	Q9P888	GIBFU	Q9p888 gibberella
39	44	45.4	764	2	Q4WKM0	ASPFU	Q4wkm0 aspergillus
40	44	45.4	1017	2	Q6VG40	SIVCZ	Q6vg40 chimpanzee
41	44	45.4	1019	1	POL_SIVS4		P12502 simian immu
42	44	45.4	1019	2	P89I54	SIVCZ	P89i54 chimpanzee
43	44	45.4	1019	2	Q7ZBR7	SIVCZ	Q7zbr7 chimpanzee
44	44	45.4	1019	2	Q7ZBR5	SIVCZ	Q7zbr5 chimpanzee
45	44	45.4	1058	1	POL_HV2D2		P15833 human immu
46	44	45.4	1387	2	Q5BFN8	EMENI	Q5bf8 aspergillus
47	44	45.4	1928	2	Q8D674	VIBVU	Q8d674 vibrio vuln
48	43.5	44.8	561	2	Q6MDL6	PARUW	Q6mdl6 parachlamyd
49	43.5	44.8	657	2	Q88L07	PSEPK	Q88l07 pseudomonas
50	43	44.3	63	2	Q98183	MCV1	Q98i83 molluscum c
51	43	44.3	113	2	Q8VLU9	9BACT	Q8vlu9 uncultured
52	43	44.3	124	2	Q48532	LEPDI	Q48532 leptothrix
53	43	44.3	215	2	Q6AM22	DESPS	Q6am22 desulfotale
54	43	44.3	216	1	PSAL	SPIOL	Q41385 spinacia ol
55	43	44.3	217	2	Q871I5	SIVCZ	Q87i15 chimpanzee
56	43	44.3	219	2	Q54C69	DICDI	Q54c69 dictyosteli
57	43	44.3	228	2	Q60T23	CAEBR	Q60t23 caenorhabdi
58	43	44.3	233	2	Q9KC77	BACHD	Q9kc77 bacillus ha
59	43	44.3	237	1	PYRF	GLOVI	Q7nk22 gloeobacter
60	43	44.3	238	2	Q8XS8	RALSO	Q8xs8 ralstonia s
61	43	44.3	241	1	YM08	PARTE	P15609 paramecium
62	43	44.3	264	2	Q4L1V5	9ENTR	Q4l1v5 salmonella
63	43	44.3	269	1	RIBF	MYCPN	P75587 mycoplasma
64	43	44.3	304	2	Q5LT69	SILPO	Q5lt69 silicibacte
65	43	44.3	340	2	Q8UN04	SIVCZ	Q8un04 chimpanzee
66	43	44.3	340	2	Q8UN03	SIVCZ	Q8un03 chimpanzee
67	43	44.3	348	2	Q51LJ5	MAGGR	Q51lj5 magnaporthe
68	43	44.3	349	2	Q7SKK9	9HIV2	Q7skk9 human immu
69	43	44.3	349	2	Q7SKK8	9HIV2	Q7skk8 human immu
70	43	44.3	352	1	ID12	PYRAE	Q8zyf6 pyrobaculum
71	43	44.3	369	2	Q830B7	ENTFA	Q830b7 enterococcu
72	43	44.3	396	2	Q90PU1	SIVCZ	Q90pu1 chimpanzee
73	43	44.3	450	2	Q9XH69	PRUAR	Q9xh69 prunus arme
74	43	44.3	454	2	Q897U6	CLOTE	Q897u6 clostridium
75	43	44.3	472	2	Q9WHZ9	9HIV2	Q9whz9 human immu
76	43	44.3	527	1	TOP6B	PYRAE	Q8zvm0 pyrobaculum
77	43	44.3	536	2	Q8VYW2	ARATH	Q8vyw2 arabidopsis
78	43	44.3	537	2	Q946D4	ARATH	Q946d4 arabidopsis
79	43	44.3	609	2	Q856X8	9CAUD	Q856x8 mycobacteri
80	43	44.3	900	2	Q8GHS7	PSERE	Q8ghs7 pseudomonas
81	43	44.3	986	2	Q57059	SIVCZ	Q57059 chimpanzee
82	43	44.3	1005	2	Q6Y8X5	9HIV1	Q6y8x5 human immu
83	43	44.3	1022	1	POL_SIVSP		P19505 simian immu
84	43	44.3	1022	2	Q903I7	SIVCZ	Q903i7 chimpanzee
85	43	44.3	1022	2	Q87956	SIVCZ	Q87956 chimpanzee
86	43	44.3	1022	2	Q87965	SIVCZ	Q87965 chimpanzee
87	43	44.3	1022	2	Q89620	SIVCZ	Q89620 chimpanzee
88	43	44.3	1022	2	Q88135	SIVCZ	Q88135 chimpanzee
89	43	44.3	1023	2	Q5AYH8	EMENI	Q5ayh8 aspergillus
90	43	44.3	1056	2	Q04097	SIVCZ	Q04097 chimpanzee
91	43	44.3	1057	1	POL_SIVAI		Q02836 simian immu
92	43	44.3	1150	2	P90246	9RETR	P90246 feline immu
93	43	44.3	1150	2	Q64F60	9RETR	Q64f60 feline immu
94	43	44.3	1157	2	Q527F0	MAGGR	Q527f0 magnaporthe
95	43	44.3	1226	2	Q6H0K6	9HIV2	Q6h0k6 human immu
96	43	44.3	1583	1	MIS4	SCHPO	Q09725 schizosacch
97	43	44.3	1896	1	VITI1	PERAM	Q9u8m0 periplaneta
98	42.5	43.8	88	2	Q6YIS5	ORYSA	Q6ys15 oryza sativ
99	42.5	43.8	275	2	Q8ZFX2	YERPE	Q8zfx2 yersinia pe
100	42.5	43.8	275	2	Q66C47	YERPS	Q66c47 yersinia ps
101	42.5	43.8	303	2	Q4JCG1	SULAC	Q4jcg1 sulfolobus
102	42.5	43.8	307	2	Q761R3	SULAC	Q761r3 sulfolobus
103	42.5	43.8	595	2	Q6NI69	CORDI	Q6ni69 corynebacte
104	42.5	43.8	2281	2	Q6RKL2	GIBMO	Q6rk12 gibberella

Q4nfi4	arthrobacte
Q8g7m6	bifidobacte
Q53mp5	oryza sativ
Q82px5	streptomyce
Q4ped7	ustilago ma
Q6akh3	desulfotale
Q9p888	gibberella
Q4wkm0	aspergillus
Q6vg40	chimpanzee
P12502	simian immu
P89i54	chimpanzee
Q7zbr7	chimpanzee
Q7zbr5	chimpanzee
P15833	human immu
Q5bf8	aspergillus
Q8d674	vibrio vuln
Q6mdl6	parachlamyd
Q88l07	pseudomonas
Q98i83	molluscum c
Q8vlu9	uncultured
Q48532	leptothrix
Q6am22	desulfotale
Q41385	spinacia ol
Q87i15	chimpanzee
Q54c69	dictyosteli
Q60t23	caenorhabdi
Q9kc77	bacillus ha
Q7nk22	gloeobacter
Q8xs8	ralstonia s
P15609	paramecium
Q4l1v5	salmonella
P75587	mycoplasma
Q5lt69	silicibacte
Q8un04	chimpanzee
Q8un03	chimpanzee
Q51lj5	magnaporthe
Q7skk9	human immu
Q7skk8	human immu
Q8zyf6	pyrobaculum
Q830b7	enterococcu
Q90pu1	chimpanzee
Q9xh69	prunus arme
Q897u6	clostridium
Q9whz9	human immu
Q8zvm0	pyrobaculum
Q8vyw2	arabidopsis
Q946d4	arabidopsis
Q856x8	mycobacteri
Q8ghs7	pseudomonas
Q57059	chimpanzee
Q6y8x5	human immu
P19505	simian immu
Q903i7	chimpanzee
Q87956	chimpanzee
Q87965	chimpanzee
Q89620	chimpanzee
Q88135	chimpanzee
Q5ayh8	aspergillus
Q04097	chimpanzee
Q02836	simian immu
P90246	feline immu
Q64f60	feline immu
Q527f0	magnaporthe
Q6h0k6	human immu
Q09725	schizosacch
Q9u8m0	periplaneta
Q6ys15	oryza sativ
Q8zfx2	yersinia pe
Q66c47	yersinia ps
Q4jcg1	sulfolobus
Q761r3	sulfolobus
Q6ni69	corynebacte
Q6rk12	gibberella

105	42	43.3	98	2	Q5P854_AZOSE	Q5p854 azoarcus sp
106	42	43.3	114	2	Q58304_PYRHO	O58304 pyrococcus
107	42	43.3	173	2	Q8D9X7_VIBVU	Q8d9x7 vibrio vuln
108	42	43.3	174	2	Q4VP81_9GAMM	Q4vp81 wolbachia p
109	42	43.3	189	2	Q9MFE1_BETVU	Q9mfe1 beta vulgar
110	42	43.3	226	2	Q8KF28_CHLTE	Q8kf28 chlorobium
111	42	43.3	238	2	Q835J7_ENTFA	Q835j7 enterococcu
112	42	43.3	244	2	Q8PPV5_XANAC	Q8ppv5 xanthomonas
113	42	43.3	266	2	Q81D70_BACCR	Q81d70 bacillus ce
114	42	43.3	296	2	Q7V459_PROMM	Q7v459 prochloroco
115	42	43.3	315	2	Q6HI35_BACHK	Q6hi35 bacillus th
116	42	43.3	319	2	Q9RKM5_STRCO	Q9rkm5 streptomyce
117	42	43.3	361	2	Q4WGN3_ASPFU	Q4wgn3 aspergillus
118	42	43.3	364	2	Q5E0T4_VIBF1	Q5e0t4 vibrio fisc
119	42	43.3	379	2	Q8PRC6_XANAC	Q8prc6 xanthomonas
120	42	43.3	386	1	ETR1_CANTR	Q8wzm3 candida tro
121	42	43.3	386	1	ETR2_CANTR	Q8wzm4 candida tro
122	42	43.3	391	2	Q6UUH2_ORYSA	Q6uuh2 oryza sativ
123	42	43.3	416	1	HEM1_AERPE	Q9y9j2 aeropyrum p
124	42	43.3	438	2	Q9MFE2_BETVU	Q9mfe2 beta vulgar
125	42	43.3	443	2	Q89FY2_BRAJA	Q89fy2 bradyrhizob
126	42	43.3	454	2	Q35213_OENBE	Q35213 oenothera b
127	42	43.3	462	2	Q4LYF0_9BURK	Q4lyf0 burkholderi
128	42	43.3	500	2	Q61FS2_CAEBR	Q61fs2 caenorhabdi
129	42	43.3	519	2	Q9SRY9_ARATH	Q9sry9 arabidopsis
130	42	43.3	580	2	Q89RH2_BRAJA	Q89rh2 bradyrhizob
131	42	43.3	581	2	Q43856_VICFA	Q43856 vicia faba
132	42	43.3	612	2	Q830J7_ENTFA	Q830j7 enterococcu
133	42	43.3	617	1	SYP_SYNY3	P73942 synecocyst
134	42	43.3	628	2	Q6BET2_CABEL	Q6bet2 caenorhabdi
135	42	43.3	635	2	Q19186_CABEL	Q19186 caenorhabdi
136	42	43.3	658	2	Q9IAC1_BRARE	Q9iac1 brachydanio
137	42	43.3	711	2	Q4UPE6_XANCP	Q4upe6 xanthomonas
138	42	43.3	711	2	Q8P3W9_XANCP	Q8p3w9 xanthomonas
139	42	43.3	711	2	Q8PFF2_XANAC	Q8pff2 xanthomonas
140	42	43.3	716	2	Q6CXN1_KLJULA	Q6cxn1 kluyveromyc
141	42	43.3	746	2	Q4LNE9_9BURK	Q4lne9 burkholderi
142	42	43.3	776	2	Q6K683_ORYSA	Q6k683 oryza sativ
143	42	43.3	805	2	Q9L538_XANOR	Q9l538 xanthomonas
144	42	43.3	822	2	Q33831_THEMA	Q33831 thermotoga
145	42	43.3	848	2	Q4RVV3_TETNG	Q4rvv3 tetraodon n
146	42	43.3	863	2	Q9ST50_MAIZE	Q9st50 zea mays (m
147	42	43.3	1015	2	Q8JAH1_SIVCZ	Q8jah1 chimpanzee
148	42	43.3	1047	1	POL_SIVAI	P27973 simian immu
149	42	43.3	1049	2	Q9XBP6_MYXXA	Q9xbp6 myxococcus
150	42	43.3	1061	1	POL_SIVAT	P05895 simian immu
151	42	43.3	1241	2	Q81TW1_BACAN	Q81tw1 bacillus an
152	42	43.3	1370	2	Q4W7J1_TOBAC	Q4w7j1 nicotiana t
153	42	43.3	1407	2	Q4T188_TETNG	Q4t188 tetraodon n
154	42	43.3	1471	2	Q54I73_DICDI	Q54i73 dictyosteli
155	42	43.3	2108	2	Q7VEZ3_MYCBO	Q7vez3 mycobacteri
156	42	43.3	2108	2	O53901_MYCTU	O53901 mycobacteri
157	42	43.3	2282	2	Q8K4S1_MOUSE	Q8k4s1 mus musculu
158	41.5	42.8	168	2	Q9V492_DROME	Q9v492 drosophila
159	41.5	42.8	308	1	XERC_CORGL	Q8mnz9 corynebacte
160	41.5	42.8	406	2	Q65EX2_BACLD	Q65ex2 bacillus li
161	41.5	42.8	877	2	O15644_TOXGO	O15644 toxoplasma
162	41	42.3	53	2	O5YWJ5_NOCPA	O5ywj5 nocardia fa
163	41	42.3	61	2	Q8VJ58_MYCTU	Q8vj58 mycobacteri
164	41	42.3	70	1	RK28_CYPAPA	P48129 cyanophora
165	41	42.3	116	2	Q6Z3C8_ORYSA	Q6z3c8 oryza sativ
166	41	42.3	135	2	Q6FSL6_CANGA	Q6fsl6 candida gla
167	41	42.3	153	2	Q7CZf6_AGRt5	Q7czf6 agrobacteri
168	41	42.3	173	2	Q5WHI4_BACSK	Q5whi4 bacillus cl
169	41	42.3	178	2	Q8UFP3_AGRt5	Q8ufp3 agrobacteri
170	41	42.3	188	2	Q9PHI6_XYLFA	Q9phi6 xylella fas
171	41	42.3	192	2	Q8SVU0_ENCCU	Q8svu0 encephalito
172	41	42.3	203	2	P87305_SCHPO	P87305 schizosacch
173	41	42.3	237	1	PYRF_ENTFA	Q834e3 enterococcu
174	41	42.3	238	2	P74728_SYNY3	P74728 synecocyst
175	41	42.3	241	2	Q6F8B0_ACIAD	Q6f8b0 acinetobact
176	41	42.3	245	2	O66272_9SPHN	O66272 erythro bact
177	41	42.3	249	2	O82989_9SPHN	O82989 erythro bact

41	42.3	251	2	Q5HLW9_STAEP	Q5hlw9 staphylococ
41	42.3	251	2	Q8CRE9_STAEP	Q8cre9 staphylococ
41	42.3	256	2	Q6GEF3_STAAR	Q6gef3 staphylococ
41	42.3	257	1	TRMD_AQUAE	O67463 aquifex ao
41	42.3	264	2	Q8YUM2_ANASP	Q8yum2 anabaena sp
41	42.3	266	2	O53002_ECOLI	O53002 escherichia
41	42.3	266	2	Q65FV4_BACLD	Q65fv4 bacillus li
41	42.3	271	2	Q4MVQ5_BACCE	Q4mvq5 bacillus ce
41	42.3	272	2	Q72XP5_BACC1	Q72xp5 bacillus ce
41	42.3	278	2	Q9XDV0_9SPHN	Q9xdv0 erythro bact
41	42.3	288	2	Q7V391_PROMP	Q7v391 prochloroco
41	42.3	293	1	ARGB_SYNPX	Q7u5c7 synechococc
41	42.3	305	2	Q6A7C7_PROAC	Q6a7c7 propionibac
41	42.3	306	2	Q7D906_MYCTU	Q7d906 mycobacteri
41	42.3	306	2	O05576_MYCTU	O05576 mycobacteri
41	42.3	306	2	Q7U0W3_MYCBO	Q7u0w3 mycobacteri
41	42.3	325	2	Q4ZSC3_PSESY	Q4zsc3 pseudomonas
41	42.3	328	2	Q6NEV8_CORDI	Q6nev8 corynebacte
41	42.3	331	2	Q4JU09_CORJK	Q4ju09 corynebacte
41	42.3	333	2	Q7MRR5_WOLSU	Q7mrr5 wolinnella s
41	42.3	338	2	Q4LPN5_9BURK	Q4lpn5 burkholderi
41	42.3	345	2	O5XH15_XENLA	Q5xh15 xenopus lae
41	42.3	369	2	O5NTF8_9BACT	Q5ntf8 uncultured
41	42.3	375	2	Q6QW16_AZOB	Q6qw16 azospirillu
41	42.3	397	1	TRPB1_ARCFU	O28672 archaeoglob
41	42.3	403	2	Q88NU2_PSEPK	Q88nu2 pseudomonas
41	42.3	419	1	CBPA1_HUMAN	P15085 homo sapien
41	42.3	419	2	O53XU0_HUMAN	Q53xu0 homo sapien
41	42.3	440	2	O5XL13_AERHY	Q5xl13 aeromonas h
41	42.3	450	2	Q9I938_XENLA	Q9i938 xenopus lae
41	42.3	450	2	Q6DFF3_XENLA	Q6dff3 xenopus lae
41	42.3	450	2	Q6P2X7_XENTR	Q6p2x7 xenopus tro
41	42.3	462	2	Q8TMW0_METAC	Q8tmw0 methanosarc
41	42.3	469	2	Q8MMK2_DUGJA	Q8mmk2 dugesia jap
41	42.3	472	2	Q7U0G1_MYCBO	Q7u0g1 mycobacteri
41	42.3	472	2	O50438_MYCTU	O50438 mycobacteri
41	42.3	474	2	Q754V5_ASHGO	Q754v5 ashbya goss
41	42.3	476	2	Q4WL94_ASPFU	Q4wl94 aspergillus
41	42.3	491	2	O5L085_GEOKA	Q5l085 geobacillus
41	42.3	496	2	Q6CER7_YARLI	Q6cer7 yarrowia li
41	42.3	503	2	Q4WDN9_ASPFU	Q4wdn9 aspergillus
41	42.3	519	2	Q4WM15_ASPFU	Q4wm15 aspergillus
41	42.3	552	2	Q83932_9RETR	Q83932 ovine lenti
41	42.3	552	2	Q83933_9RETR	Q83933 ovine lenti
41	42.3	552	2	Q83934_9RETR	Q83934 ovine lenti
41	42.3	552	2	Q83935_9RETR	Q83935 ovine lenti
41	42.3	576	2	Q90892_CHICK	Q90892 gallus gall
41	42.3	590	2	Q4IH39_GIBZE	Q4ih39 gibberella
41	42.3	594	2	Q5FRH3_GLUOX	Q5frh3 gluconobact
41	42.3	607	2	Q87HD5_VIBPA	Q87hd5 vibrio para
41	42.3	622	1	MTL1_LACLA	P35516 lactococcus
41	42.3	664	2	Q6LKN3_PHOPR	Q6lkn3 photobacter
41	42.3	679	2	O5IU94_MAGGR	O5iu94 magnaporth
41	42.3	695	2	Q4RH59_TETNG	Q4rh59 tetraodon n
41	42.3	715	2	Q6PAA9_XENLA	Q6paa9 xenopus lae
41	42.3	721	2	Q8K6D6_STRP3	Q8k6d6 streptococc
41	42.3	721	2	Q5XAM7_STRP6	Q5xam7 streptococc
41	42.3	721	2	Q99YL1_STRPY	Q99yl1 streptococc
41	42.3	721	2	Q7CMY6_STRP8	Q7cmy6 streptococc
41	42.3	801	2	O5AS91_EMENI	O5as91 aspergillus
41	42.3	810	2	O5TFJ4_HUMAN	O5tfj4 homo sapien
41	42.3	889	2	Q9FI79_STRGR	Q9fi79 streptomyce
41	42.3	897	2	Q8EI13_SHEON	Q8ei13 shewanella
41	42.3	967	2	Q4FP31_9RICK	Q4fp31 candidatus
41	42.3	1007	2	Q699W7_SIVCZ	Q699w7 chimpanzee
41	42.3	1007	2	Q699V9_SIVCZ	Q699v9 chimpanzee
41	42.3	1018	2	Q9IBU9_9HIV2	Q9ib19 human immu
41	42.3	1039	2	Q87108_SIVCZ	Q87108 chimpanzee
41	42.3	1046	1	POL_SIVAG	P27980 simian immu
41	42.3	1055	2	Q89928_9HIV2	Q89928 human immu
41	42.3	1060	2	O5QGQ1_SIVCZ	O5ggq1 chimpanzee
41	42.3	1060	2	O5QFT5_SIVCZ	O5qft5 chimpanzee
41	42.3	1060	2	O5QFR3_SIVCZ	O5qfr3 chimpanzee

178	41	42.3	251	2	Q5HLW9_STAEP	Q5hlw9 staphylococ
179	41	42.3	251	2	Q8CRE9_STAEP	Q8cre9 staphylococ
180	41	42.3	256	2	Q6GEF3_STAAR	Q6gef3 staphylococ
181	41	42.3	257	1	TRMD_AQUAE	O67463 aquifex ao
182	41	42.3	264	2	Q8YUM2_ANASP	Q8yum2 anabaena sp
183	41	42.3	266	2	O53002_ECOLI	O53002 escherichia
184	41	42.3	266	2	Q65FV4_BACLD	Q65fv4 bacillus li
185	41	42.3	271	2	Q4MVQ5_BACCE	Q4mvq5 bacillus ce
186	41	42.3	272	2	Q72XP5_BACC1	Q72xp5 bacillus ce
187	41	42.3	278	2	Q9XDV0_9SPHN	Q9xdv0 erythro bact
188	41	42.3	288	2	Q7V391_PROMP	Q7v391 prochloroco
189	41	42.3	293	1	ARGB_SYNPX	Q7u5c7 synechococc
190	41	42.3	305	2	Q6A7C7_PROAC	Q6a7c7 propionibac
191	41	42.3	306	2	Q7D906_MYCTU	Q7d906 mycobacteri
192	41	42.3	306	2	O05576_MYCTU	O05576 mycobacteri
193	41	42.3	306	2	Q7U0W3_MYCBO	Q7u0w3 mycobacteri
194	41	42.3	325	2	Q4ZSC3_PSESY	Q4zsc3 pseudomonas
195	41	42.3	328	2	Q6NEV8_CORDI	Q6nev8 corynebacte
196	41	42.3	331	2	Q4JU09_CORJK	Q4ju09 corynebacte
197	41	42.3	333	2	Q7MRR5_WOLSU	Q7mrr5 wolinnella s
198	41	42.3	338	2	Q4LPN5_9BURK	Q4lpn5 burkholderi
199	41	42.3	345	2	O5XH15_XENLA	Q5xh15 xenopus lae
200	41	42.3	369	2	O5NTF8_9BACT	Q5ntf8 uncultured
201	41	42.3	375	2	Q6QW16_AZOB	Q6qw16 azospirillu
202	41	42.3	397	1	TRPB1_ARCFU	O28672 archaeoglob
203	41	42.3	403	2	Q88NU2_PSEPK	Q88nu2 pseudomonas
204	41	42.3	419	1	CBPA1_HUMAN	P15085 homo sapien
205	41	42.3	419	2	O53XU0_HUMAN	Q53xu0 homo sapien
206	41	42.3	440	2	O5XL13_AERHY	Q5xl13 aeromonas h
207	41	42.3	450	2	Q9I938_XENLA	Q9i938 xenopus lae
208	41	42.3	450	2	Q6DFF3_XENLA	Q6dff3 xenopus lae
209	41	42.3	450	2	Q6P2X7_XENTR	Q6p2x7 xenopus tro
210	41	42.3	462	2	Q8TMW0_METAC	Q8tmw0 methanosarc
211	41	42.3	469	2	Q8MMK2_DUGJA	Q8mmk2 dugesia jap
212	41	42.3	472	2	Q7U0G1_MYCBO	Q7u0g1 mycobacteri
213	41	42.3	472	2	O50438_MYCTU	O50438 mycobacteri
214	41	42.3	474	2	Q754V5_ASHGO	Q754v5 ashbya goss
215	41	42.3	476	2	Q4WL94_ASPFU	Q4wl94 aspergillus
216	41	42.3	491	2	O5L085_GEOKA	Q5l085 geobacillus</

251	41	42.3	1060	2	Q5QFL2_SIVCZ	Q5qfl2 chimpanzee	324	40	41.2	349	2	Q9RQL6_STRPN	Q9rql6 streptococc
252	41	42.3	1060	2	Q5QFJ6_SIVCZ	Q5qfj6 chimpanzee	325	40	41.2	349	2	Q4JYG2_STRPN	Q4jyg2 streptococc
253	41	42.3	1061	2	Q9W699_FUGRU	Q9w699 fugu rubrip	326	40	41.2	349	2	Q4JYN4_STRPN	Q4jyn4 streptococc
254	41	42.3	1061	2	Q76634_9HIV2	Q76634 human immun	327	40	41.2	349	2	Q4JYW5_STRPN	Q4jyw5 streptococc
255	41	42.3	1064	2	Q8UTV7_9HIV2	Q8utv7 human immun	328	40	41.2	349	2	Q4JYV6_STRPN	Q4jyv6 streptococc
256	41	42.3	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis	329	40	41.2	349	2	Q4K034_STRPN	Q4k034 streptococc
257	41	42.3	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis	330	40	41.2	349	2	Q4K0A7_STRPN	Q4k0a7 streptococc
258	41	42.3	1124	1	POL_FIVSD	P19028 f pol polyp	331	40	41.2	349	2	Q4K0Q5_STRPN	Q4k0q5 streptococc
259	41	42.3	1140	1	Y376_MYCPN	P75405 mycoplasma	332	40	41.2	349	2	Q4K0W1_STRPN	Q4k0w1 streptococc
260	41	42.3	1191	2	O52484_SERMA	O52484 serrattia ma	333	40	41.2	349	2	Q4K121_STRPN	Q4k121 streptococc
261	41	42.3	1820	2	Q54F55_DICDI	Q54f55 dictyosteli	334	40	41.2	349	2	Q4K2P4_STRPN	Q4k2p4 streptococc
262	41	42.3	3433	1	PCX_DRÖME	P18490 adoxophila	335	40	41.2	349	2	Q4K2T4_STRPN	Q4k2t4 streptococc
263	40.5	41.8	324	2	Q7T9Q0_GVAC	Q7t9q0 adoxophyes	336	40	41.2	349	2	O54546_STRPN	O54546 streptococc
264	40.5	41.8	504	2	Q6EMC3_ECOLI	Q6emc3 escherichia	337	40	41.2	349	2	Q8DR64_STRR6	Q8dr64 streptococc
265	40.5	41.8	504	2	Q9WTH8_9ZZZZ	Q9wth8 plasmid r10	338	40	41.2	349	2	Q7SKK6_9HIV2	Q7skk6 human immun
266	40.5	41.8	507	1	TRA6_PSEAE	P0a134 pseudomonas	339	40	41.2	349	2	Q7SKK4_9HIV2	Q7skk4 human immun
267	40.5	41.8	507	1	TRA6_SHIFL	P0a135 shigella fl	340	40	41.2	349	2	Q7SKJ5_9HIV2	Q7skj5 human immun
268	40.5	41.8	507	2	Q7BT57_ECOLI	Q7bt57 escherichia	341	40	41.2	349	2	Q7SKJ4_9HIV2	Q7skj4 human immun
269	40.5	41.8	507	2	Q79PC0_KLEPN	Q79pc0 klebsiella	342	40	41.2	349	2	Q7SKJ3_9HIV2	Q7skj3 human immun
270	40.5	41.8	507	2	Q79CH1_PSEAE	Q79ch1 pseudomonas	343	40	41.2	349	2	Q7SKJ2_9HIV2	Q7skj2 human immun
271	40.5	41.8	889	2	Q4IUH0_AZOV1	Q4iuh0 azotobacter	344	40	41.2	349	2	Q7SKJ1_9HIV2	Q7skj1 human immun
272	40	41.2	81	2	Q9NDL7_HYDMA	Q9ndl7 hydra magni	345	40	41.2	349	2	Q7SKJ0_9HIV2	Q7skj0 human immun
273	40	41.2	100	2	Q6VFT3_9POXV	Q6vft3 squirrel po	346	40	41.2	349	2	Q7SKI8_9HIV2	Q7ski8 human immun
274	40	41.2	108	2	Q51C64_ENTHI	Q5lc64 entamoeba h	347	40	41.2	349	2	Q7SKI7_9HIV2	Q7ski7 human immun
275	40	41.2	115	2	Q4PN77_IXOSC	Q4pn77 ixodes scap	348	40	41.2	349	2	Q7SKI6_9HIV2	Q7ski6 human immun
276	40	41.2	129	1	RS9_WOLSU	Q7m7r1 wolinnella s	349	40	41.2	349	2	Q7SKI5_9HIV2	Q7ski5 human immun
277	40	41.2	129	2	Q8DHX7_SYNEL	Q8dhx7 synechococc	350	40	41.2	349	2	Q7SKI4_9HIV2	Q7ski4 human immun
278	40	41.2	133	2	Q4ZPV7_PSESY	Q4zpv7 pseudomonas	351	40	41.2	349	2	Q7SKI3_9HIV2	Q7ski3 human immun
279	40	41.2	134	2	Q88615_PSESM	Q88615 pseudomonas	352	40	41.2	349	2	Q7SKI2_9HIV2	Q7ski2 human immun
280	40	41.2	139	2	Q9GSF9_HYDAT	Q9gsf9 hydra atten	353	40	41.2	349	2	Q7SKI1_9HIV2	Q7ski1 human immun
281	40	41.2	146	2	Q6WMX3_ANAPH	Q6wmx3 anaplasma p	354	40	41.2	351	2	Q9E3T1_9HIV2	Q9e3t1 human immun
282	40	41.2	154	2	Q9WI01_9HIV2	Q9wi01 human immun	355	40	41.2	351	2	Q9E3T2_9HIV2	Q9e3t2 human immun
283	40	41.2	158	2	Q9X0A5_THEMEA	Q9x0a5 thermotoga	356	40	41.2	351	2	Q9E3T4_9HIV2	Q9e3t4 human immun
284	40	41.2	165	2	Q5PX28_9CHLO	Q5px28 halimeda tu	357	40	41.2	351	2	TSN10_HUMAN	Q9hiz9 homo sapien
285	40	41.2	165	2	Q5RI40_BRARE	Q5ri40 brachydanio	358	40	41.2	355	1	Q8NKN4_9CREN	Q8nkn4 uncultured
286	40	41.2	172	2	Q9G8T3_9CRYP	Q9g8t3 rhodomonas	359	40	41.2	361	2	O53605_STAAU	O53605 staphylococ
287	40	41.2	182	2	Q9P5V8_NEUCR	Q9p5v8 neurospora	360	40	41.2	368	2	Q69LD7_ORYSA	Q69ld7 oryza sativ
288	40	41.2	185	2	Q7RVZ3_NEUCR	Q7rvz3 neurospora	361	40	41.2	369	2	Q6AGP0_LEIXX	Q6agp0 leifsonia x
289	40	41.2	193	1	PTPM1_MOUSE	Q66gt5 mus musculu	362	40	41.2	374	2	Q6SEA7_9CAUD	Q6sea7 lactobacill
290	40	41.2	207	1	GST3_CAEEL	O161l6 caenorhabdi	363	40	41.2	391	2	Q65PW8_LACUO	Q65pw8 lactobacill
291	40	41.2	211	2	P88796_9HIV1	P88796 human immun	364	40	41.2	391	2	Q56919_9HIV1	Q56919 human immun
292	40	41.2	212	2	Q4KLJ2_STRPN	Q4klj2 streptococc	365	40	41.2	394	2	O56920_9HIV1	O56920 human immun
293	40	41.2	236	2	Q9XVJ3_CAEEL	Q9xvj3 caenorhabdi	366	40	41.2	394	2	O56917_9HIV1	O56917 human immun
294	40	41.2	242	1	Y1085_MYCTU	P67157 mycobacteri	367	40	41.2	395	1	2AAA_PEA	P36875 pisum sativ
295	40	41.2	242	1	Y1114_MYCBO	P67158 mycobacteri	368	40	41.2	396	2	Q7ZMY2_9HIV1	Q7zmy2 human immun
296	40	41.2	244	2	Q6ZN98_HUMAN	Q6zn98 homo sapien	369	40	41.2	399	2	Q8RID4_FUSNN	Q8rid4 fusobacteri
297	40	41.2	249	2	Q9A535_CAUCR	Q9a535 caulobacter	370	40	41.2	401	2	Q7VVG2_BORPE	Q7vvg2 bordetella
298	40	41.2	251	2	Q4KAL0_PSEF5	Q4kal0 pseudomonas	371	40	41.2	401	2	Q7WB96_BORPE	Q7wb96 bordetella
299	40	41.2	261	2	Q5IR77_9HIV1	Q5ir77 human immun	372	40	41.2	401	2	Q7WMR4_BORBR	Q7wmr4 bordetella
300	40	41.2	261	2	Q5IR73_9HIV1	Q5ir73 human immun	373	40	41.2	403	2	Q8KDE7_CHLTE	Q8kde7 chlorobium
301	40	41.2	262	2	Q9HWM4_PSEAE	Q9hwm4 pseudomonas	374	40	41.2	412	2	O8IN73_DROME	O8in73 drosophila
302	40	41.2	265	2	O33643_SALGL	Q33643 salmonella	375	40	41.2	419	1	CBPA1_MOUSE	Q7tpz8 mus musculu
303	40	41.2	283	2	Q6ANH6_DESPS	Q6anh6 desulfotale	376	40	41.2	419	1	CBPA1_RAT	P00731 rattus norv
304	40	41.2	288	2	Q7U3V2_SYNPX	Q7u3v2 synechococc	377	40	41.2	419	2	Q9D4K7_MOUSE	Q9d4k7 mus musculu
305	40	41.2	297	2	Q7UQE4_RHOBA	Q7uqe4 rhodopirell	378	40	41.2	428	2	Q38614_BPTU2	Q38614 lactococcus
306	40	41.2	306	2	Q9Z5G1_MYCLE	Q9z5g1 mycobacteri	379	40	41.2	428	2	MTA1_ACEPA	O52702 acetobacter
307	40	41.2	310	2	Q7SKI9_9HIV2	Q7ski9 human immun	380	40	41.2	429	2	Q37969_9CAUD	Q37969 bacterioph
308	40	41.2	312	2	Q978P0_THEVO	Q978p0 thermoplasm	381	40	41.2	429	2	Q8LTP4_9CAUD	Q8ltp4 lactococcus
309	40	41.2	325	2	Q5QYT8_IDILO	Q5qyt8 idiomarina	382	40	41.2	429	2	Q9AZ52_9CAUD	Q9az52 lactococcus
310	40	41.2	330	2	Q84M50_ORYSA	Q84m50 oryza sativ	383	40	41.2	429	2	Q9FZS8_9CAUD	Q9fzs8 bacterioph
311	40	41.2	336	2	Q9WYF8_THEMEA	Q9wyf8 thermotoga	384	40	41.2	429	2	Q9FTZ3_9CAUD	Q9ftz3 bacterioph
312	40	41.2	344	2	Q9HN36_HALSA	Q9hn36 halobacteri	385	40	41.2	439	2	Q98700_SIVCZ	Q98700 chimpanzee
313	40	41.2	348	2	Q6L5P7_STRMT	Q6l5p7 streptococc	386	40	41.2	443	2	Q7R3Z8_GIALA	Q7r3z8 giardia lam
314	40	41.2	348	2	Q6L5R8_STROR	Q6l5r8 streptococc	387	40	41.2	457	2	Q7ULE0_RHOBA	Q7ule0 rhodopirell
315	40	41.2	348	2	Q83YT4_STRGN	Q83yt4 streptococc	388	40	41.2	469	2	Q7XBC8_MAIZE	Q7xbc8 zea mays (m
316	40	41.2	348	2	Q9Z641_STRPN	Q9z641 streptococc	389	40	41.2	469	2	Q8AAL8_BACTN	Q8aal8 bacteroides
317	40	41.2	348	2	Q87SS3_VIBPA	Q87ss3 vibrio para	390	40	41.2	476	2	Q9WI00_9HIV2	Q9wi00 human immun
318	40	41.2	349	2	Q50Z69_ENTHI	Q50z69 entamoeba h	391	40	41.2	488	2	Q93383_CAEEL	Q93383 caenorhabdi
319	40	41.2	349	2	O54611_STRPN	O54611 streptococc	392	40	41.2	489	2	O62100_CAEEL	O62100 caenorhabdi
320	40	41.2	349	2	P97006_STRPN	P97006 streptococc	393	40	41.2	495	2	Q4WR57_ASPFU	Q4wr57 aspergillus
321	40	41.2	349	2	O08436_STRPN	O08436 streptococc	394	40	41.2	497	2	Q5CSN2_CRYPV	Q5csn2 cryptospori
322	40	41.2	349	2	Q9AH82_STRPN	Q9ah82 streptococc	395	40	41.2	497	2	Q5CJ09_CRYHO	Q5cj09 cryptospori
323	40	41.2	349	2	Q9AHAB_STRPN	Q9aha8 streptococc	396	40	41.2				

397	40	41.2	500	2	Q4V7B5 RAT	Q4v7b5 rattus norv
398	40	41.2	520	2	Q4IE36_GIBZE	Q4ie36 gibberella
399	40	41.2	521	2	O82507_ARATH	O82507 arabidopsis
400	40	41.2	525	2	Q6CCQ9_YARLI	Q6ccq9 yarrowia li
401	40	41.2	527	2	Q25564_HELPY	Q25564 helicobacte
402	40	41.2	535	2	Q5NE19_LYCES	Q5ne19 lycopersico
403	40	41.2	540	2	Q6ALB3_DESPS	Q6alb3 desulfotale
404	40	41.2	543	2	O94185_ASPOR	O94185 aspergillus
405	40	41.2	553	2	Q7WGH7_BORBR	Q7wgh7 bordetella
406	40	41.2	553	2	Q7W512_BORPA	Q7w512 bordetella
407	40	41.2	553	2	Q7VWX5_BORPE	Q7vwx5 bordetella
408	40	41.2	556	2	Q6ONX9_CAEBR	Q6onx9 caenorhabdi
409	40	41.2	560	2	Q06347_9HIV2	Q06347 human immun
410	40	41.2	565	2	Q4IRR7_GIBZE	Q4irr7 gibberella
411	40	41.2	583	2	Q7YTY5_DROME	Q7yty5 drosophila
412	40	41.2	584	2	Q57XR0_9TRYP	Q57xr0 trypanosoma
413	40	41.2	586	2	Q828W4_STRAW	Q828w4 streptomyce
414	40	41.2	601	2	Q8MSS3_DROME	Q8mss3 drosophila
415	40	41.2	601	2	Q9W3R0_DROME	Q9w3r0 drosophila
416	40	41.2	612	1	AMYG_ASPOR	P36914 aspergillus
417	40	41.2	622	2	Q6BML3_DEBHA	Q6bml3 debaryomyce
418	40	41.2	623	1	ARE1_SACBA	Q876l3 saccharomyc
419	40	41.2	666	2	Q72Y70_BACC1	Q72y70 bacillus ce
420	40	41.2	666	2	Q81X11_BACAN	Q81xi1 bacillus an
421	40	41.2	666	2	Q6HBP2_BACHK	Q6hbp2 bacillus th
422	40	41.2	678	2	Q6GNS6_XENLA	Q6gns6 xenopus lae
423	40	41.2	688	2	Q9FZR2_9CAUD	Q9fzr2 mycoplasma
424	40	41.2	722	2	Q6LD23_ERWCT	Q6ld2d3 erwinia car
425	40	41.2	753	2	Q4LTC0_9BURK	Q4ltc0 burkholderi
426	40	41.2	772	1	NUDT3_ARATH	Q81831 arabidopsis
427	40	41.2	819	2	Q5H1C4_XANOR	Q5h1c4 xanthomonas
428	40	41.2	822	2	Q6H6R9_ORYSA	Q6h6r9 oryza sativ
429	40	41.2	832	2	Q5FAN0_APIME	Q5fan0 apis mellif
430	40	41.2	898	2	Q9W1S3_DROME	Q9w1s3 drosophila
431	40	41.2	933	2	Q677U1_9VIRU	Q677u1 lymphocysti
432	40	41.2	936	2	Q7ZJ29_SIVCZ	Q7zj29 simian immu
433	40	41.2	941	2	Q5BCP6_EMENI	Q5bcp6 aspergillus
434	40	41.2	978	2	Q6G892_STAAS	Q6g892 staphylococ
435	40	41.2	978	2	Q6GFL3_STAAR	Q6gfl3 staphylococ
436	40	41.2	978	2	Q5HET2_STAAC	Q5het2 staphylococ
437	40	41.2	978	2	Q7A0K3_STAAN	Q7a0k3 staphylococ
438	40	41.2	978	2	Q7A4V5_STAAN	Q7a4v5 staphylococ
439	40	41.2	978	2	Q99T35_STAAM	Q99t35 staphylococ
440	40	41.2	1005	2	Q6Y8U8_9HIV1	Q6y8u8 human immun
441	40	41.2	1007	2	Q66QF6_9HIV1	Q66qf6 human immun
442	40	41.2	1009	1	POL_SIVGB	P22382 simian immu
443	40	41.2	1009	2	Q6Y929_9HIV1	Q6y929 human immun
444	40	41.2	1012	2	Q6Y8U0_9HIV1	Q6y8u0 human immun
445	40	41.2	1016	2	Q8W0Y3_MAIZE	Q8w0y3 zea mays (m
446	40	41.2	1027	2	Q9WPP5_SIVCZ	Q9wpp5 chimpanzee
447	40	41.2	1033	2	Q79666_9HIV1	Q79666 human immun
448	40	41.2	1034	1	POL_HV2CA	P24107 human immun
449	40	41.2	1035	2	Q73194_9HIV2	Q73194 human immun
450	40	41.2	1035	2	Q6R793_9HIV2	Q6r793 human immun
451	40	41.2	1035	2	Q6R784_9HIV2	Q6r784 human immun
452	40	41.2	1036	1	POL_HV2RO	P04584 human immun
453	40	41.2	1040	2	Q9T014_ARATH	Q9t014 arabidopsis
454	40	41.2	1043	2	Q90DD2_SIVCZ	Q90dd2 chimpanzee
455	40	41.2	1049	1	POL_HV2G1	P18042 human immun
456	40	41.2	1054	1	POL_SIVMK	P05897 simian immu
457	40	41.2	1054	2	Q5TYK3_SIVCZ	Q5tyk3 chimpanzee
458	40	41.2	1055	1	POL_HV2ST	P20876 human immun
459	40	41.2	1055	2	Q9YTU0_9HIV2	Q9ytu0 human immun
460	40	41.2	1056	1	POL_SIVM1	P05896 simian immu
461	40	41.2	1059	2	O56317_9PLVG	O56317 simian-huma
462	40	41.2	1059	2	O11404_9PLVG	O11404 simian-huma
463	40	41.2	1059	2	O56320_9PLVG	O56320 simian-huma
464	40	41.2	1059	2	Q76630_9HIV2	Q76630 human immun
465	40	41.2	1059	2	O56613_9PLVG	O56613 simian-huma
466	40	41.2	1059	2	Q5XLN1_9PLVG	Q5xln1 simian-huma
467	40	41.2	1060	2	Q90EX3_SIVCZ	Q90ex3 simian immu
468	40	41.2	1060	2	Q87706_SIVCZ	Q87706 chimpanzee
469	40	41.2	1060	2	Q88016_SIVCZ	Q88016 chimpanzee

Q90ex8 simian immu	2	Q90EX8_SIVCZ	1060	41.2	470
Q8usp8 simian-huma	2	Q8USP8_9PLVG	1060	41.2	471
Q07387 chimpanzee	2	Q07387_SIVCZ	1060	41.2	472
Q5w9h4 chimpanzee	2	Q5W9H4_SIVCZ	1060	41.2	473
Q5qgp0 chimpanzee	2	Q5QGp0_SIVCZ	1060	41.2	474
Q5qgh9 chimpanzee	2	Q5QGH9_SIVCZ	1060	41.2	475
Q5qgh0 chimpanzee	2	Q5QGH0_SIVCZ	1060	41.2	476
Q5qgf7 chimpanzee	2	Q5QGF7_SIVCZ	1060	41.2	477
Q5qgb6 chimpanzee	2	Q5QGB6_SIVCZ	1060	41.2	478
Q5qga9 chimpanzee	2	Q5QGA9_SIVCZ	1060	41.2	479
Q5qg99 chimpanzee	2	Q5QG99_SIVCZ	1060	41.2	480
Q5qg92 chimpanzee	2	Q5QG92_SIVCZ	1060	41.2	481
Q5qg28 chimpanzee	2	Q5QG28_SIVCZ	1060	41.2	482
Q5qfx8 chimpanzee	2	Q5QFX8_SIVCZ	1060	41.2	483
Q5qfv7 chimpanzee	2	Q5QFV7_SIVCZ	1060	41.2	484
Q5qfu2 chimpanzee	2	Q5QFU2_SIVCZ	1060	41.2	485
Q5qfs2 chimpanzee	2	Q5QFS2_SIVCZ	1060	41.2	486
Q5qfq4 chimpanzee	2	Q5QFQ4_SIVCZ	1060	41.2	487
Q5qfp4 chimpanzee	2	Q5QFP4_SIVCZ	1060	41.2	488
Q5qfn4 chimpanzee	2	Q5QFN4_SIVCZ	1060	41.2	489
Q5qfm3 chimpanzee	2	Q5QFM3_SIVCZ	1060	41.2	490
Q5qfk8 chimpanzee	2	Q5QFK8_SIVCZ	1060	41.2	491
Q5qfi9 chimpanzee	2	Q5QFI9_SIVCZ	1060	41.2	492
Q5izc2 chimpanzee	2	Q5IZC2_SIVCZ	1060	41.2	493
Q4xzk5 plasmodium	2	Q4XZK5_PLACH	1090	41.2	494
Q9mb43 adiantum ca	2	Q9MB43_ADICA	1092	41.2	495
Q9i7s4 drosophila	2	Q9I7S4_DROME	1105	41.2	496
Q4p945 ustilago ma	2	Q4P945_USTMA	1110	41.2	497
Q66933 feline immu	2	Q66933_9RETR	1123	41.2	498
Q6c802 yarrowia li	2	Q6C802_YARLI	1139	41.2	499
Q54yu8 dictyosteli	2	Q54YU8_DICDI	1210	41.2	500
Q69u98 oryza sativ	2	Q69U98_ORYSA	1485	41.2	501
Q4n585 theileria p	2	Q4N585_THEPA	1552	41.2	502
Q7x7y2 oryza sativ	2	Q7X7Y2_ORYSA	1629	41.2	503
Q5vza9 homo sapien	2	Q5VZA9_HUMAN	1706	41.2	504
P08572 homo sapien	1	CO4A2_HUMAN	1712	41.2	505
P34367 caenorhabdi	1	YLJ2_CAEEL	2107	41.2	506
O74390 schizosacch	2	O74390_SCHPO	2609	41.2	507
Q5wpu4 usutu virus	2	Q5WPU4_9FLAV	3434	41.2	508
Q5wpu5 usutu virus	2	Q5WPU5_9FLAV	3434	41.2	509
Q6w5q0 streptomyce	2	Q6W5Q0_9ACTO	10625	41.2	510
Q9fgp7 arabidopais	2	Q9FGP7_ARATH	202	40.7	511
Q5wld4 bacillus cl	2	Q5WLD4_BACSK	202	40.7	512
Q5bg06 aspergillus	2	Q5BG06_EMENI	318	40.7	513
Q4q5m2 leishmania	2	Q4Q5M2_LEIMA	363	40.7	514
Q52ds4 magnaporth	2	Q52DS4_MAGGR	431	40.7	515
Q9vp06 drosophila	2	Q9VP06_DROME	475	40.7	516
Q4idq2 gibberella	2	Q4IDQ2_GIBZE	580	40.7	517
Q4zr95 pseudomonas	2	Q4ZR95_PSESY	642	40.7	518
Q9hzi6 pseudomonas	2	Q9HZI6_PSEAE	642	40.7	519
Q606u4 methylococc	2	Q606U4_METCA	718	40.7	520
O13839 schizosacch	1	FIN1_SCHPO	722	40.7	521
Q6xge6 escherichia	2	Q6XGE6_ECOLI	964	40.7	522
Q22830 caenorhabdi	2	Q22830_CAEEL	1746	40.7	523
Q51xm7 magnaporth	2	Q51XM7_MAGGR	4643	40.7	524
Q74ed1 geobacter s	2	Q74ED1_GEOSL	46	40.2	525
Q605u7 methylococc	2	Q605U7_METCA	70	40.2	526
Q9ndl5 tima formos	2	Q9NDL5_9CNID	81	40.2	527
Q9ndl8 hydractinia	2	Q9NDL8_HYDEC	81	40.2	528
Q9ndl9 eirene sp.	2	Q9NDL9_9CNID	81	40.2	529
Q67pb8 symbiobacte	2	Q67PB8_SYMTH	98	40.2	530
Q82r21 streptomyce	2	Q82R21_STRAW	98	40.2	531
Q8bsd1 mus musculu	2	Q8BSD1_MOUSE	108	40.2	532
Q8n9n4 homo sapien	2	Q8N9N4_HUMAN	126	40.2	533
Q6wgq4 gloeophyllu	2	Q6WGO4_GLOTR	136	40.2	534
Q5zfk0 populus tre	2	Q5ZFK0_POPTN	136	40.2	535
Q6d8e8 erwinia car	2	Q6D8E8_ERWCT	151	40.2	536
Q54xf9 dictyosteli	2	Q54XF9_DICDI	154	40.2	537
Q5px26 halimeda tu	2	Q5PX26_9CHLO	165	40.2	538
Q4gx23 halimeda hu	2	Q4GX23_9CHLO	165	40.2	539
Q96zs9 sulfolobus	2	Q96ZS9_SULTO	174	40.2	540
Q4vp36 francisella	2	Q4VP36_FRATT	174	40.2	541
Q4vp38 francisella	2	Q4VP38_FRATT	174	40.2	542

543	39	40.2	174	2	Q4VP41_FRATU	Q4vp41 francisella	616	39	40.2	294	2	Q94FC9_MAIZE	Q94fc9 zea mays (m
544	39	40.2	174	2	Q4VP45_FRANO	Q4vp45 francisella	617	39	40.2	294	2	Q94FD0_MAIZE	Q94fd0 zea mays (m
545	39	40.2	174	2	Q4VP47_FRATU	Q4vp47 francisella	618	39	40.2	296	2	Q6R4X2_ZEAMP	Q6r4x2 zea mays (s
546	39	40.2	174	2	Q4VP40_FRATU	Q4vp40 francisella	619	39	40.2	296	2	Q6R4X4_ZEAMP	Q6r4x4 zea mays (s
547	39	40.2	183	2	Q9LUE5_ARATH	Q9lue5 arabidopsis	620	39	40.2	296	2	Q6R4Y1_ZEAMM	Q6r4y1 zea mays (s
548	39	40.2	188	2	Q5Z7Q8_ORYSA	Q5z7q8 oryza sativ	621	39	40.2	296	2	Q6R4Y3_ZEAMM	Q6r4y3 zea mays (s
549	39	40.2	195	2	Q9S2G3_STRCO	Q9s2g3 streptomyce	622	39	40.2	296	2	Q6R4Y4_ZEALU	Q6r4y4 zea luxuria
550	39	40.2	205	1	COAE_GLOVI	Q7nlp9 gloeobacter	623	39	40.2	296	2	Q6R4Y5_ZEALU	Q6r4y5 zea luxuria
551	39	40.2	205	2	Q55LP5_CRYNE	Q55lp5 cryptococcu	624	39	40.2	296	2	Q6R4Y6_ZEALU	Q6r4y6 zea luxuria
552	39	40.2	206	2	Q5NGD3_FRATT	Q5ngd3 francisella	625	39	40.2	296	2	Q6R4Y7_ZEALU	Q6r4y7 zea luxuria
553	39	40.2	206	2	Q6MEP0_PARUW	Q6mep0 parachlamyd	626	39	40.2	296	2	Q6R4Y8_ZEALU	Q6r4y8 zea luxuria
554	39	40.2	212	2	Q4JU70_CORJK	Q4ju70 corynebacte	627	39	40.2	296	2	Q6R4Y9_ZEALU	Q6r4y9 zea luxuria
555	39	40.2	214	2	Q7X3Y6_VIBVU	Q7x3y6 vibrio vuln	628	39	40.2	296	2	Q6R4Z1_ZEALU	Q6r4z1 zea luxuria
556	39	40.2	214	2	Q5LGQ5_BACFN	Q5lqg5 bacteroides	629	39	40.2	296	2	Q93W75_MAIZE	Q93w75 zea mays (m
557	39	40.2	214	2	Q6F7K2_ACIAD	Q6f7k2 acinetobact	630	39	40.2	296	2	Q93W76_MAIZE	Q93w76 zea mays (m
558	39	40.2	214	2	Q64XK3_BACFR	Q64xk3 bacteroides	631	39	40.2	296	2	Q94FC7_MAIZE	Q94fc7 zea mays (m
559	39	40.2	215	2	Q6D962_ERWCT	Q6d962 erwinia car	632	39	40.2	296	2	Q94FC8_MAIZE	Q94fc8 zea mays (m
560	39	40.2	217	2	Q89ZQ0_BACTN	Q89zq0 bacteroides	633	39	40.2	298	2	Q6R4W7_ZEAMP	Q6r4w7 zea mays (s
561	39	40.2	218	2	Q526F1_MAGGR	Q526f1 magnaporth	634	39	40.2	298	2	Q6R4W8_ZEAMP	Q6r4w8 zea mays (s
562	39	40.2	220	1	NADD_YERPE	Q8zdg1 yersinia pe	635	39	40.2	298	2	Q6R4X0_ZEAMP	Q6r4x0 zea mays (s
563	39	40.2	220	2	Q66DE6_YERPS	Q66de6 yersinia ps	636	39	40.2	298	2	Q6R4X1_ZEAMP	Q6r4x1 zea mays (s
564	39	40.2	227	2	Q645E4_9SALA	Q645e4 oedipina po	637	39	40.2	298	2	Q6R4X7_ZEAMM	Q6r4x7 zea mays (s
565	39	40.2	234	2	Q7N956_PHOLL	Q7n956 photorhabd	638	39	40.2	298	2	Q6R4X9_ZEAMM	Q6r4x9 zea mays (s
566	39	40.2	238	2	Q12158_9HIV1	Q12158 human immun	639	39	40.2	298	2	Q6R4Y0_ZEAMM	Q6r4y0 zea mays (s
567	39	40.2	240	1	HIS4_LISMF	Q722y6 listeria mo	640	39	40.2	298	2	Q53QG3_ORYSA	Q53qg3 oryza sativ
568	39	40.2	240	1	HIS4_LISMO	Q8y9g4 listeria mo	641	39	40.2	300	2	Q6R4X5_ZEAMM	Q6r4x5 zea mays (s
569	39	40.2	240	1	UNG_MYCPN	P75536 mycoplasma	642	39	40.2	301	2	Q64WZ2_BACFR	Q64wz2 bacteroides
570	39	40.2	241	2	Q6UU23_ORYSA	Q6uu23 oryza sativ	643	39	40.2	301	2	Q82GR2_STRAW	Q82gr2 streptomyce
571	39	40.2	253	2	Q52575_PSESX	Q52575 pseudomonas	644	39	40.2	301	2	Q810P1_MOUSE	Q810p1 mus musculu
572	39	40.2	256	1	VMAT_ORSVW	Q84l31 ovine respi	645	39	40.2	304	2	Q82S45_NITEU	Q82s45 nitrosomona
573	39	40.2	256	2	Q77L00_BRSPA	Q77l00 bovine resp	646	39	40.2	311	2	Q7ZYD7_XENLA	Q7zyd7 xenopus lae
574	39	40.2	256	2	Q9YS27_9MONO	Q9ys27 bovine resp	647	39	40.2	314	1	LEUO_ECOLI	P10151 escherichia
575	39	40.2	268	2	Q6T5X5_9POAL	Q6t5x5 bouteloua h	648	39	40.2	314	2	Q7AHR5_ECO57	Q7ahr5 escherichia
576	39	40.2	268	2	Q6T5X6_9POAL	Q6t5x6 bouteloua h	649	39	40.2	314	2	Q89D85_BRAJA	Q89d85 bradyrhizob
577	39	40.2	268	2	Q6T5X7_9POAL	Q6t5x7 bouteloua h	650	39	40.2	314	2	Q8FL73_ECOL6	Q8fl73 escherichia
578	39	40.2	268	2	Q6T5X8_9POAL	Q6t5x8 bouteloua h	651	39	40.2	318	1	NIA_CHLVU	Q01170 chlorella v
579	39	40.2	268	2	Q6T5X9_9POAL	Q6t5x9 bouteloua h	652	39	40.2	319	2	Q84NR4_ORYSA	Q84nr4 oryza sativ
580	39	40.2	268	2	Q6T5Y0_9POAL	Q6t5y0 bouteloua h	653	39	40.2	320	2	Q8X9Z7_ECO57	Q8x9z7 escherichia
581	39	40.2	268	2	Q6T5Y2_9POAL	Q6t5y2 bouteloua h	654	39	40.2	324	2	Q75QN0_9VIRU	Q75qn0 bacterioph
582	39	40.2	268	2	Q6T5Y3_9POAL	Q6t5y3 bouteloua h	655	39	40.2	325	2	Q68334_VIBCH	Q68334 vibrio chol
583	39	40.2	268	2	Q6T5Y4_9POAL	Q6t5y4 bouteloua h	656	39	40.2	325	2	Q8D4Y4_VIBVU	Q8d4y4 vibrio vuln
584	39	40.2	268	2	Q6T5Y5_9POAL	Q6t5y5 bouteloua h	657	39	40.2	327	2	Q9KTS1_VIBCH	Q9kts1 vibrio chol
585	39	40.2	268	2	Q6T5Y6_9POAL	Q6t5y6 bouteloua h	658	39	40.2	327	2	Q9D9S9_MOUSE	Q9d9s9 mus musculu
586	39	40.2	268	2	Q6T5Y7_9POAL	Q6t5y7 bouteloua h	659	39	40.2	329	2	Q9DZC1_9HIV1	Q9dzc1 human immun
587	39	40.2	268	2	Q6T5Y8_9POAL	Q6t5y8 bouteloua h	660	39	40.2	329	2	Q9DZC2_9HIV1	Q9dzc2 human immun
588	39	40.2	268	2	Q6T5Y0_9POAL	Q6t5y0 bouteloua h	661	39	40.2	330	1	G3PC_LEIME	Q01558 leishmania
589	39	40.2	268	2	Q6T5Z4_9POAL	Q6t5z4 bouteloua h	662	39	40.2	331	2	P93795_TRIDA	P93795 tripsacum d
590	39	40.2	268	2	Q6TAD1_9POAL	Q6tad1 bouteloua t	663	39	40.2	336	1	ASB1_MOUSE	Q9wv74 mus musculu
591	39	40.2	268	2	Q6TAD7_9POAL	Q6tad7 bouteloua d	664	39	40.2	336	1	TS2_MAIZE	P50160 zea mays (m
592	39	40.2	268	2	Q6TAD8_9POAL	Q6tad8 bouteloua d	665	39	40.2	338	2	Q4P1Z3_USTMA	Q4p1z3 ustilago ma
593	39	40.2	268	2	Q6TAD9_9POAL	Q6tad9 bouteloua d	666	39	40.2	341	2	Q58YZ9_LACRH	Q58yz9 lactobacill
594	39	40.2	268	2	Q6TAE1_9POAL	Q6tae1 bouteloua d	667	39	40.2	341	2	Q58Z16_LACRH	Q58z16 lactobacill
595	39	40.2	268	2	Q6TAE2_9POAL	Q6tae2 bouteloua d	668	39	40.2	341	2	Q58Z33_LACRH	Q58z33 lactobacill
596	39	40.2	268	2	Q6TAE3_9POAL	Q6tae3 bouteloua d	669	39	40.2	342	2	Q88083_ENTFA	Q88083 enterococcu
597	39	40.2	268	2	Q6TAE4_9POAL	Q6tae4 bouteloua d	670	39	40.2	342	2	Q5SIS7_THET8	Q5sis7 thermus the
598	39	40.2	268	2	Q6TAE5_9POAL	Q6tae5 bouteloua d	671	39	40.2	342	2	Q72J58_THET2	Q72j58 thermus the
599	39	40.2	268	2	Q6TAE7_9POAL	Q6tae7 bouteloua d	672	39	40.2	342	2	Q832N1_ENTFA	Q832n1 enterococcu
600	39	40.2	268	2	Q6TAE8_9POAL	Q6tae8 bouteloua d	673	39	40.2	342	2	Q80TI9_MOUSE	Q80ti9 mus musculu
601	39	40.2	268	2	Q6TAE9_9POAL	Q6tae9 bouteloua d	674	39	40.2	349	2	Q9XD92_ECOLI	Q9xd92 escherichia
602	39	40.2	268	2	Q6TAF0_9POAL	Q6taf0 bouteloua d	675	39	40.2	349	2	Q9CJ08_ECOLI	Q9cj08 lactococcus
603	39	40.2	268	2	Q6TAF3_9POAL	Q6taf3 bouteloua d	676	39	40.2	349	2	Q5R0S2_IDILO	Q5r0s2 idiomarina
604	39	40.2	268	2	Q6TAF4_9POAL	Q6taf4 bouteloua d	677	39	40.2	349	2	Q8GGW5_LACLC	Q8ggw5 lactococcus
605	39	40.2	268	2	Q6TAF5_9POAL	Q6taf5 bouteloua d	678	39	40.2	355	2	Q9RJU4_STRCO	Q9rju4 streptococ
606	39	40.2	273	2	Q67LW2_SYMTH	Q67lw2 symbiobacte	679	39	40.2	356	2	Q842N5_VIBCH	Q842n5 vibrio chol
607	39	40.2	274	2	Q5WA99_ORYSA	Q5wa99 oryza sativ	680	39	40.2	359	2	Q5FT89_GLUOX	Q5ft89 gluconobact
608	39	40.2	276	2	Q6R4X8_ZEAMM	Q6r4x8 zea mays (s	681	39	40.2	363	2	Q5P236_AJECA	Q5p236 ajelomyces
609	39	40.2	279	2	Q4HYE6_GIBZE	Q4hye6 gibberella	682	39	40.2	364	1	MRAW_BORBR	Q7wfr5 bordetella
610	39	40.2	287	2	Q5N4Z4_SYNP6	Q5n4z4 synechococc	683	39	40.2	364	1	MRAW_BORPA	Q7w4a7 bordetella
611	39	40.2	289	2	Q6R4W5_ZEAMP	Q6r4w5 zea mays (s	684	39	40.2	364	1	MRAW_BORPE	Q7vup6 bordetella
612	39	40.2	291	2	Q8PHT6_XANAC	Q8pht6 xanthomonas	685	39	40.2	365	2	Q6C009_YARLI	Q6c009 yarrowia li
613	39	40.2	294	2	Q93VX4_MAIZE	Q93vx4 zea mays (m	686	39	40.2	366	2	Q6CDN8_YARLI	Q6cdn8 yarrowia li
614	39	40.2	294	2	Q93VX5_MAIZE	Q93vx5 zea mays (m	687	39	40.2	368	2	Q6C1S2_YARLI	Q6c1s2 yarrowia li
615	39	40.2	294	2	Q93W09_MAIZE	Q93w09 zea mays (m	688	39	40.2	368	2	Q6C5R4_YARLI	Q6c5r4 yarrowia li

689	39	40.2	368	2	Q6CSX6_YARLI	Q6c5x6 yarrowia li
690	39	40.2	368	2	Q6CBy8_YARLI	Q6cby8 yarrowia li
691	39	40.2	368	2	Q6CCN0_YARLI	Q6ccn0 yarrowia li
692	39	40.2	368	2	Q6CHS0_YARLI	Q6ch50 yarrowia li
693	39	40.2	369	2	Q877S5_BRAJA	Q877s5 bradyrhizob
694	39	40.2	369	2	Q89YD1_BRAJA	Q89yd1 bradyrhizob
695	39	40.2	371	2	Q6CDZ5_YARLI	Q6cdz5 yarrowia li
696	39	40.2	372	2	Q5WFP2_BACSK	Q5wfp2 bacillus cl
697	39	40.2	373	2	Q83SN9_SHIFL	Q83sn9 shigella fl
698	39	40.2	381	2	Q7V9U3_PROWA	Q7v9u3 prochloroc
699	39	40.2	395	2	Q83XI6_VIBAN	Q83xi6 vibrio angu
700	39	40.2	396	2	Q51EK9_ENTHI	Q51ek9 entamoeba h
701	39	40.2	403	2	Q7Y0B1_ORYSA	Q7y0b1 oryza sativ
702	39	40.2	404	2	Q6NAS6_RHOPA	Q6nas6 rhodopseu
703	39	40.2	405	2	Q5PZ38_AJECA	Q5pz38 ajellomyces
704	39	40.2	410	2	Q9LMR4_ARATH	Q9lmr4 arabidopsis
705	39	40.2	414	2	Q4L061_9HIV1	Q4l061 human immu
706	39	40.2	426	2	Q6SGV9_9BACT	Q6sgv9 uncultured
707	39	40.2	427	2	Q6KZ55_PICTO	Q6kz55 microphilus
708	39	40.2	429	1	SYN_THEAC	Q9hks7 thermoplasm
709	39	40.2	429	1	SYN_THEVO	Q979y4 thermoplasm
710	39	40.2	429	2	Q7S2Z7_NEUCR	Q7s2z7 neurospora
711	39	40.2	438	2	Q5M9U2_TOBAC	Q5m9u2 nicotiana t
712	39	40.2	441	2	Q87HJ3_VIBPA	Q87hj3 vibrio para
713	39	40.2	442	2	Q6W4T1_VIBAN	Q6w4t1 vibrio angu
714	39	40.2	445	2	Q70WP6_AERSA	Q70wp6 aeromonas s
715	39	40.2	448	2	Q6CMI5_KLULA	Q6cmi5 kluyveromyc
716	39	40.2	462	2	Q5BP96_AEDAE	Q5bp96 aedes aegyp
717	39	40.2	465	2	Q6ZFM2_ORYSA	Q6zfm2 oryza sativ
718	39	40.2	471	2	Q7W4T3_BORPA	Q7w4t3 bordetella
719	39	40.2	473	2	Q7PKB4_ANOGA	Q7pkb4 anopheles g
720	39	40.2	474	2	Q7VU97_BORPE	Q7vu97 bordetella
721	39	40.2	474	2	Q7WGA9_BORBR	Q7wga9 bordetella
722	39	40.2	475	2	Q6KAQ6_MOUSE	Q6kaq6 mus musculu
723	39	40.2	481	2	Q4P4B5_USTMA	Q4p4b5 ustilago ma
724	39	40.2	486	2	Q6EY2_YERPS	Q6evy2 yersinia ps
725	39	40.2	489	1	OCLN_POTTR	Q28793 potorous tr
726	39	40.2	491	2	Q5V4L8_HALMA	Q5v4l8 haloarcula
727	39	40.2	492	2	Q8USW7_9HIV1	Q8usw7 human immu
728	39	40.2	504	2	Q8GW59_ARATH	Q8gw59 arabidopsis
729	39	40.2	508	2	Q6JSN2_9MYRI	Q6jsn2 sphaerother
730	39	40.2	516	2	Q5CW26_CRYPV	Q5cw26 cryptospori
731	39	40.2	516	2	Q5CG37_CRYHO	Q5cg37 cryptospori
732	39	40.2	519	2	Q5XIS3_RAT	Q5xi83 rattus norv
733	39	40.2	531	2	Q7U8T8_SYNPX	Q7u8t8 synechococc
734	39	40.2	533	2	Q9KGJ0_BACHD	Q9kgj0 bacillus ha
735	39	40.2	537	2	Q91F86_IRV6	Q91f86 chilo iride
736	39	40.2	539	2	Q5UYP5_HALMA	Q5uyp5 haloarcula
737	39	40.2	550	2	Q5B1T8_EMENI	Q5b1t8 aspergillus
738	39	40.2	551	2	Q5Z8N9_ORYSA	Q5z8n9 oryza sativ
739	39	40.2	551	2	Q8SB30_ORYSA	Q8sb30 oryza sativ
740	39	40.2	554	2	Q870V7_NEUCR	Q870v7 neurospora
741	39	40.2	562	2	Q9W409_DROME	Q9w409 drosophila
742	39	40.2	567	2	Q27090_TRIVA	Q27090 trichomonas
743	39	40.2	567	2	Q27093_TRIVA	Q27093 trichomonas
744	39	40.2	571	1	DNA43_YEAST	P32354 saccharomyc
745	39	40.2	577	2	Q63X25_BURPS	Q63x25 burkholderi
746	39	40.2	580	2	Q82DF0_STRAW	Q82df0 streptomyce
747	39	40.2	584	2	Q84DD9_VIBPA	Q84dd9 vibrio para
748	39	40.2	584	2	Q87FY4_VIBPA	Q87fy4 vibrio para
749	39	40.2	591	2	Q5XIT5_RAT	Q5xit5 rattus norv
750	39	40.2	592	2	Q4ICT8_GIBZE	Q4ict8 gibberella
751	39	40.2	607	2	Q73HG5_WOLPM	Q73hg5 wolbachia p
752	39	40.2	609	2	Q9JMM7_9RICK	Q9jmm7 wolbachia s
753	39	40.2	609	2	Q60CU6_METCA	Q60cu6 methylococc
754	39	40.2	610	2	Q9JMP6_9VIRU	Q9jmp6 bacterioph
755	39	40.2	619	2	Q4KKX0_HUMAN	Q4kkx0 homo sapien
756	39	40.2	619	2	Q69JZ1_ORYSA	Q69jz1 oryza sativ
757	39	40.2	644	1	VE1_HPVS8	P26543 human papil
758	39	40.2	658	2	Q9BNW5_9HEXA	Q9bnw5 tomocerus s
759	39	40.2	664	2	Q8D3P3_VIBVU	Q8d3p3 vibrio vuln
760	39	40.2	664	2	Q7MF63_VIBVY	Q7mf63 vibrio vuln
761	39	40.2	690	1	EPB42_HUMAN	P16452 homo sapien

762	39	40.2	691	2	Q4VB97_HUMAN	Q4vb97 homo sapien
763	39	40.2	720	2	Q663C4_YERPS	Q663c4 yersinia ps
764	39	40.2	721	1	ENP1_TORCA	P14400 torpedo cal
765	39	40.2	728	2	Q6JUC0_9MYRI	Q6juc0 abacion mag
766	39	40.2	731	2	Q6BM04_DEBHA	Q6bm04 debaryomyce
767	39	40.2	741	2	Q5NHO1_FRATT	Q5nhq1 francisella
768	39	40.2	753	2	Q7XQR1_ORYSA	Q7xqr1 oryza sativ
769	39	40.2	754	2	Q9SUI2_ARATH	Q9sui2 arabidopsis
770	39	40.2	778	2	Q6UUI4_ORYSA	Q6uui4 oryza sativ
771	39	40.2	817	2	Q7QE40_ANOGA	Q7qe40 anopheles g
772	39	40.2	818	2	Q6PBA6_BRARE	Q6pba6 brachydanio
773	39	40.2	836	2	Q9W0J2_DROME	Q9w0j2 drosophila
774	39	40.2	859	2	Q04095_ARATH	Q04095 arabidopsis
775	39	40.2	877	2	Q42497_CHLVU	Q42497 chlorella v
776	39	40.2	899	2	Q4ZT52_PSESY	Q4zt52 pseudomonas
777	39	40.2	899	2	Q881L7_PSESM	Q881l7 pseudomonas
778	39	40.2	937	2	Q7UN41_RHOBA	Q7un41 rhodopirell
779	39	40.2	946	2	Q9VJU4_DROME	Q9vj04 drosophila
780	39	40.2	948	2	Q8S5C7_ORYSA	Q8s5c7 oryza sativ
781	39	40.2	965	1	PSD1_CAEEL	Q18115 caenorhabdi
782	39	40.2	970	2	Q60KI7_CAEBR	Q60k17 caenorhabdi
783	39	40.2	971	2	Q52G37_MAGGR	Q52g37 magnaporth
784	39	40.2	980	2	Q8T8Q8_DROME	Q8t8q8 drosophila
785	39	40.2	980	2	Q9W495_DROME	Q9w495 drosophila
786	39	40.2	981	2	Q8ASV8_BACTN	Q8asv8 bacteroides
787	39	40.2	1004	2	Q8CGA7_MOUSE	Q8cga7 mus musculu
788	39	40.2	1006	2	Q4Q7I3_LEIMA	Q4q7i3 leishmania
789	39	40.2	1016	2	Q8JAI0_SIVCZ	Q8jai0 chimpanzee
790	39	40.2	1034	2	Q8VHL7_MOUSE	Q8vhl7 mus musculu
791	39	40.2	1034	2	Q8VIK5_MOUSE	Q8vik5 mus musculu
792	39	40.2	1046	2	Q7PVU9_ANOGA	Q7pvu9 anopheles g
793	39	40.2	1049	2	Q88099_SIVCZ	Q88099 chimpanzee
794	39	40.2	1050	1	LKTA_ACTAC	P16462 actinobacil
795	39	40.2	1055	2	Q43892_ACTAC	Q43892 actinobacil
796	39	40.2	1063	2	Q7UP09_RHOBA	Q7up09 rhodopirell
797	39	40.2	1067	2	P89904_SIVCZ	P89904 chimpanzee
798	39	40.2	1122	2	Q8JAV1_9RETR	Q8j4v1 visna/maedi
799	39	40.2	1164	2	Q561B0_CRYNE	Q561b0 cryptococcu
800	39	40.2	1168	2	Q4IPF9_GIBZE	Q4ipf9 gibberella
801	39	40.2	1185	2	Q5KQ11_CRYNE	Q5kq11 cryptococcu
802	39	40.2	1215	2	Q4P4F6_USTMA	Q4p4f6 ustilago ma
803	39	40.2	1246	2	Q560T0_CRYNE	Q560t0 cryptococcu
804	39	40.2	1246	2	Q5KPS9_CRYNE	Q5kps9 cryptococcu
805	39	40.2	1275	2	Q54Y47_DICDI	Q54y47 dictyosteli
806	39	40.2	1339	2	Q9UST7_SCHPO	Q9ust7 schizosacch
807	39	40.2	1534	2	Q8JG61_BRARE	Q8jg61 brachydanio
808	39	40.2	1544	1	CEND3_HUMAN	Q8wvn8 homo sapien
809	39	40.2	1566	2	Q7FAP7_ORYSA	Q7fap7 oryza sativ
810	39	40.2	2302	2	Q8IZA4_HUMAN	Q8iza4 homo sapien
811	39	40.2	2396	2	Q8SA91_MAIZE	Q8sa91 zea mays (m
812	39	40.2	4189	2	Q6BL44_DEBHA	Q6bl44 debaryomyce
813	39	40.2	5072	2	Q6JD66_TOXGO	Q6jd66 toxoplasma
814	38.5	39.7	137	2	Q6FEL1_ACTIAD	Q6fel1 acinetobact
815	38.5	39.7	193	1	KCY_METKA	Q8tzb3 methanopyru
816	38.5	39.7	246	2	Q9FC61_STRCO	Q9fc61 streptomyce
817	38.5	39.7	256	2	Q6G741_STAAS	Q6g741 staphylococ
818	38.5	39.7	256	2	Q5HDS6_STAAC	Q5hds6 staphylococ
819	38.5	39.7	256	2	Q7A436_STAAN	Q7a436 staphylococ
820	38.5	39.7	256	2	Q8NV94_STAAM	Q8nv94 staphylococ
821	38.5	39.7	256	2	Q99RZ1_STAAM	Q99rz1 staphylococ
822	38.5	39.7	331	2	Q63U03_BURPS	Q63u03 burkholderi
823	38.5	39.7	340	2	Q6AEX6_LEIXX	Q6aex6 leifsonia x
824	38.5	39.7	381	2	Q6ZIJ7_ORYSA	Q6zij7 oryza sativ
825	38.5	39.7	416	1	RPOD_MTCAB	P52322 microcystis
826	38.5	39.7	420	2	Q9P5Z3_NEUCR	Q9p5z3 neurospora
827	38.5	39.7	430	2	Q7Q159_ANOGA	Q7q159 anopheles g
828	38.5	39.7	435	2	Q5TP03_ANOGA	Q5tp03 anopheles g
829	38.5	39.7	439	2	Q96MU6_HUMAN	Q96mu6 homo sapien
830	38.5	39.7	455	2	Q9GSW2_DROVI	Q9gsW2 drosophila
831	38.5	39.7	479	2	Q61K83_CAEBR	Q61k83 caenorhabdi
832	38.5	39.7	481	2	Q7RUZ6_NEUCR	Q7ruz6 neurospora
833	38.5	39.7	489	2	Q6R573_9HYPO	Q6r573 trichoderma
834	38.5	39.7	494	2	Q9GSW1_DROVI	Q9gsW1 drosophila

Q6c5x6	yarrowia li
Q6cby8	yarrowia li
Q6ccn0	yarrowia li
Q6ch50	yarrowia li
Q877s5	bradyrhizob
Q89yd1	bradyrhizob
Q6cdz5	yarrowia li
Q5wfp2	bacillus cl
Q83sn9	shigella fl
Q7v9u3	prochloroc
Q83xi6	vibrio angu
Q51ek9	entamoeba h
Q7y0b1	oryza sativ
Q6nas6	rhodopseu
Q5pz38	ajellomyces
Q9lmr4	arabidopsis
Q4l061	human immu
Q6sgv9	uncultured
Q6kz55	microphilus
Q9hks7	thermoplasm
Q979y4	thermoplasm
Q7s2z7	neurospora
Q5m9u2	nicotiana t
Q87hj3	vibrio para
Q6w4t1	vibrio angu
Q70wp6	aeromonas s
Q6cmi5	kluyveromyc
Q5bp96	aedes aegyp
Q6zfm2	oryza sativ
Q7w4t3	bordetella
Q7pkb4	anopheles g
Q7vu97	bordetella
Q7wga9	bordetella
Q6kaq6	mus musculu
Q4p4b5	ustilago ma
Q6evy2	yersinia ps
Q28793	potorous tr
Q5v4l8	haloarcula
Q8usw7	human immu
Q8gw59	arabidopsis
Q6jsn2	sphaerother
Q5cw26	cryptospori
Q5cg37	cryptospori
Q5xi83	rattus norv
Q7u8t8	synechococc
Q9kgj0	bacillus ha
Q91f86	chilo iride
Q5uyp5	haloarcula
Q5b1t8	aspergillus
Q5z8n9	oryza sativ
Q8ab30	oryza sativ
Q870v7	neurospora
Q9w409	drosophila
Q27090	trichomonas
Q27093	trichomonas
P32354	saccharomyc
Q63x25	burkholderi
Q82df0	streptomyce
Q84dd9	vibrio para
Q87fy4	vibrio para
Q5xit5	rattus norv
Q4ict8	gibberella
Q73hg5	wolbachia p
Q9jmm7	wolbachia s
Q60cu6	methylococc
Q9jmp6	bacterioph
Q4kkx0	homo sapien
Q69jz1	oryza sativ
P26543	human papil
Q9bnw5	tomocerus s
Q8d3p3	vibrio vuln
Q7mf63	vibrio vuln
P16452	homo sapien

835	38.5	39.7	516	2	Q6M3H3_CORGL	Q6m3h3	corynebacte	908	38	39.2	228	2	Q9VRA4_DROME	Q9vra4	drosophila
836	38.5	39.7	519	2	Q6C3N9_YARLI	Q6c3n9	yarrowia li	909	38	39.2	228	2	Q8SZE4_DROME	Q8sze4	drosophila
837	38.5	39.7	520	2	Q8NND9_CORGL	Q8nd9	corynebacte	910	38	39.2	229	2	Q8KMU6_ENTFA	Q8kmu6	enterococcu
838	38.5	39.7	529	2	Q57IL8_SALCH	Q57il8	salmonella	911	38	39.2	229	2	Q88T17_LACPL	Q88t17	lactobacill
839	38.5	39.7	529	2	Q5PJQ0_SALPA	Q5pjq0	salmonella	912	38	39.2	230	2	Q7ZKL1_9HIV1	Q7zkl1	human immun
840	38.5	39.7	529	2	Q8Z278_SALTI	Q8z278	salmonella	913	38	39.2	241	2	Q21302_CAEEL	Q21302	caenorhabdi
841	38.5	39.7	529	2	Q8ZLC7_SALTY	Q8zlc7	salmonella	914	38	39.2	241	2	O66269_9SPHN	Q66269	erythrombri
842	38.5	39.7	529	2	Q9GSW3_DROVI	Q9gs3	drosophila	915	38	39.2	245	2	O66278_9SPHN	Q66278	agrobacteri
843	38.5	39.7	587	2	Q5CXV8_CRYPV	Q5cxv8	cryptospori	916	38	39.2	245	2	Q9ZN87_9SPHN	Q9zn87	porphyrobac
844	38.5	39.7	587	2	Q5CLN1_CRYHO	Q5cln1	cryptospori	917	38	39.2	246	2	O66276_9SPHN	Q66276	porphyrobac
845	38.5	39.7	639	2	Q7NRI4_CHRVO	Q7nri4	chromobacte	918	38	39.2	246	2	Q9I9S7_PHACC	Q9i9s7	phasianus c
846	38.5	39.7	659	2	Q4J480_AZOVI	Q4j480	azotobacter	919	38	39.2	247	2	O24659_ORYSA	Q24659	oryza sativ
847	38.5	39.7	665	2	Q870B6_9FUNG	Q870b6	piromyces s	920	38	39.2	247	2	Q8S1X5_ORYSA	Q8s1x5	oryza sativ
848	38.5	39.7	857	2	Q24736_DROVI	Q24736	drosophila	921	38	39.2	250	2	Q42648_BETVU	Q42648	beta vulgar
849	38.5	39.7	857	2	Q6S015_DROME	Q6s015	drosophila	922	38	39.2	250	2	Q88VQ0_LACPL	Q88vq0	lactobacill
850	38.5	39.7	857	2	Q6S016_DROME	Q6s016	drosophila	923	38	39.2	250	2	O11771_9PARA	Q11771	measles vir
851	38.5	39.7	857	2	Q95VB5_DROVI	Q95vb5	drosophila	924	38	39.2	250	2	O11776_9PARA	Q11776	measles vir
852	38.5	39.7	1073	1	RAG1_ONCMY	Q91187	oncorhynchu	925	38	39.2	250	2	O11777_9PARA	Q11777	measles vir
853	38.5	39.7	2476	1	ZAN_FIG	Q28983	sus scrofa	926	38	39.2	254	1	AZLC_BACSU	O07942	bacillus su
854	38	39.2	31	2	Q4YGC2_PLABE	Q4ygc2	plasmodium	927	38	39.2	256	1	VMAT_HRSVA	P03419	human respi
855	38	39.2	68	2	Q7R2G9_GIALA	Q7r2g9	giardia lam	928	38	39.2	256	1	Q77YB3_HRSV	Q77yb3	human respi
856	38	39.2	84	2	Q6XK08_SPICI	Q6xk08	spiroplasma	929	38	39.2	256	2	Q4KRW7_HRSV	Q4krw7	human respi
857	38	39.2	84	2	Q6C8D7_YARLI	Q6c8d7	yarrowia li	930	38	39.2	257	2	Q9Z948_CHLPN	Q9z948	chlamydia p
858	38	39.2	99	2	Q4HKA2_CAMLA	Q4hka2	campylobact	931	38	39.2	260	2	Q49657_MYCLE	Q49657	mycobacteri
859	38	39.2	101	2	Q5ZC53_ORYSA	Q5zc53	oryza sativ	932	38	39.2	261	2	Q5IR78_9HIV1	Q5ir78	human immun
860	38	39.2	102	2	Q5L786_CHLAB	Q5l786	chlamydophi	933	38	39.2	262	2	Q9YQ24_SIVCZ	Q9yq24	chimpanzee
861	38	39.2	104	2	Q92U58_RHIME	Q92u58	rhizobium m	934	38	39.2	278	2	Q9FY71_ARATH	Q9fy71	arabidopsis
862	38	39.2	119	2	Q7SFR6_NEUCR	Q7sfr6	neurospora	935	38	39.2	279	2	Q6IGU0_DROME	Q6igu0	drosophila
863	38	39.2	125	2	Q5QNA1_ORYSA	Q5qna1	oryza sativ	936	38	39.2	281	2	Q9SN47_ARATH	Q9sn47	arabidopsis
864	38	39.2	134	2	Q9RW17_DEIRA	Q9rw17	deinococcus	937	38	39.2	282	2	Q6ZPT4_MOUSE	Q6zpt4	mus musculu
865	38	39.2	135	2	Q4TPZ6_9SPHN	Q4tpz6	erythrobact	938	38	39.2	285	2	Q9HLI1_THEAC	Q9hli1	thermoplasm
866	38	39.2	136	2	Q9DAL0_MOUSE	Q9dal0	mus musculu	939	38	39.2	285	2	Q5YFP7_9VIRU	Q5yfp7	singapore g
867	38	39.2	140	2	Q8EUP0_MYCPE	Q8eup0	mycoplasma	940	38	39.2	286	2	Q5GAL3_9VIRU	Q5gal3	grouper iri
868	38	39.2	147	2	Q82Y81_NITEU	Q82y81	nitrosomona	941	38	39.2	286	2	Q7NWB3_CHRVO	Q7nwb3	chromobacte
869	38	39.2	151	2	Q4Q3Q5_LEIMA	Q4q3q5	leishmania	942	38	39.2	287	2	Q6LPY7_PHOPR	Q6lpy7	photobacter
870	38	39.2	153	2	Q6AMX6_DESPS	Q6amx6	desulfotale	943	38	39.2	287	2	Q7VDZ5_PROMA	Q7vdz5	prochloroco
871	38	39.2	154	2	Q7NZJ2_CHRVO	Q7nzj2	chromobacte	944	38	39.2	298	2	Q8BV58_MOUSE	Q8bv58	mus musculu
872	38	39.2	159	2	Q9AXE8_CITUN	Q9axe8	citrus unsh	945	38	39.2	301	2	Q4NHT3_9MICC	Q4nht3	arthrobacte
873	38	39.2	160	2	Q73UZ7_MYCPA	Q73uz7	mycobacteri	946	38	39.2	301	2	Q8FAB5_ECOL6	Q8fab5	escherichia
874	38	39.2	163	2	Q83JD4_SHIFL	Q83jd4	shigella fl	947	38	39.2	306	2	Q5RQF9_9HIV1	Q5rqf9	human immun
875	38	39.2	169	2	Q9ZSW7_HAMVI	Q9zsw7	hamamelis v	948	38	39.2	306	2	Q8EJ00_SHEON	Q8ej00	shewanella
876	38	39.2	171	2	Q9SBC1_HAMVI	Q9sbc1	hamamelis v	949	38	39.2	306	2	Q83NN4_TROW8	Q83nn4	tropheryma
877	38	39.2	172	2	Q7RHA0_PLAYO	Q7rha0	plasmodium	950	38	39.2	309	2	Q73H24_WOLPM	Q73h24	wolbachia p
878	38	39.2	173	2	Q9SBC0_HAMVI	Q9sbc0	hamamelis v	951	38	39.2	311	1	LYSR_ECOLI	P03030	escherichia
879	38	39.2	173	2	Q9ZSW6_HAMVI	Q9zsw6	hamamelis v	952	38	39.2	311	2	Q7AB47_ECO57	Q7ab47	escherichia
880	38	39.2	175	2	Q5U648_HUMAN	Q5u648	homo sapien	953	38	39.2	311	2	Q57KA3_SALCH	Q57ka3	salmonella
881	38	39.2	179	2	Q5KY62_GEOKA	Q5ky62	geobacillus	954	38	39.2	311	2	Q8X6J2_ECO57	Q8x6j2	escherichia
882	38	39.2	180	2	Q5V2S6_HALMA	Q5v2s6	halocaulula	955	38	39.2	311	2	Q8FEA1_ECOL6	Q8feal	escherichia
883	38	39.2	181	2	Q9F3Z6_NEIME	Q9f3z6	neisseria m	956	38	39.2	311	2	Q7N7G0_PHOLI	Q7n7g0	photorhabdu
884	38	39.2	182	2	Q92VU0_RHIME	Q92vu0	rhizobium m	957	38	39.2	311	2	Q7CPV5_SALTY	Q7cpv5	salmonella
885	38	39.2	183	2	Q4Q0C4_LEIMA	Q4q0c4	leishmania	958	38	39.2	311	2	Q5PEP7_SALPA	Q5pep7	salmonella
886	38	39.2	183	2	Q8D7J6_VIBVU	Q8d7j6	vibrio vuln	959	38	39.2	311	2	Q8XGD5_SALTI	Q8xgd5	salmonella
887	38	39.2	184	2	Q4NNK9_9DELT	Q4nnk9	anaeromyxob	960	38	39.2	311	2	Q83QB6_SHIFL	Q83qb6	shigella fl
888	38	39.2	185	2	Q975X8_SULTO	Q975x8	sulfolobus	961	38	39.2	312	2	Q5PT51_MOUSE	Q5pt51	mus musculu
889	38	39.2	190	2	Q7Q0R7_ANOGA	Q7q0r7	anopheles g	962	38	39.2	313	2	Q667E7_YERPS	Q667e7	yersinia ps
890	38	39.2	190	2	Q6DT74_ARALP	Q6dt74	arabidopsis	963	38	39.2	316	2	Q83MV4_TROWT	Q83mv4	tropheryma
891	38	39.2	191	2	Q9RJX0_STRCO	Q9rjx0	streptomyce	964	38	39.2	320	2	Q982F1_RHILO	Q982f1	rhizobium l
892	38	39.2	194	1	Y004_MYCSM	Q50434	mycobacteri	965	38	39.2	321	2	O39288_9ALPH	Q39288	equid herpe
893	38	39.2	196	2	Q8SAP5_9BRYO	Q8sap5	mitthyridiu	966	38	39.2	322	2	Q9PTB4_BRARE	Q9ptb4	brachydanio
894	38	39.2	198	1	RECR_GLUOX	Q5ftb6	gluconobact	967	38	39.2	327	1	Y745_HELPY	O25441	helicobacte
895	38	39.2	199	2	Q9X6K1_KLEOX	Q9x6k1	klebsiella	968	38	39.2	327	2	Q5H9S6_HUMAN	Q5h9s6	homo sapien
896	38	39.2	203	1	IFNB_CHICK	Q90873	gallus gall	969	38	39.2	328	2	Q8CZX8_YERPE	Q8czx8	yersinia pe
897	38	39.2	203	2	Q5MFU9_CHICK	Q5mfu9	gallus gall	970	38	39.2	331	2	Q9YA52_AERPE	Q9ya52	aeropyrum p
898	38	39.2	206	2	Q622A8_CAEBR	Q622a8	caenorhabdi	971	38	39.2	331	2	Q6NJB6_CORDI	Q6njb6	corynebacte
899	38	39.2	208	2	Q51DP1_ENTHI	Q51dp1	entamoeba h	972	38	39.2	333	2	Q5ZJH8_CHICK	Q5zjh8	gallus gall
900	38	39.2	209	2	Q5N2P0_SYNP6	Q5n2p0	synechococc	973	38	39.2	335	1	Y479_STAAN	Q8nxy9	staphylococ
901	38	39.2	210	2	Q6IGL5_DROME	Q6igl5	drosophila	974	38	39.2	335	1	Y481_STAAS	Q6gbw4	staphylococ
902	38	39.2	214	2	Q8DDE8_VIBVU	Q8dde8	vibrio vuln	975	38	39.2	335	1	Y524_STAAM	P65206	staphylococ
903	38	39.2	214	2	Q7MGK1_VIBVU	Q7mgk1	vibrio vuln	976	38	39.2	335	1	Y527_STAAR	P65205	staphylococ
904	38	39.2	222	2	Q9U796_MUSDO	Q9u796	musca domes	977	38	39.2	335	1	Y527_STAAR	Q6gje5	staphylococ
905	38	39.2	223	2	Q9XJ28_ORYSA	Q9xj28	oryza sativ	978	38	39.2	335	2	Q4JDJ2_9PROT	Q4jdj2	nitrosomona
906	38	39.2	225	2	Q7MBL3_WOLSU	Q7mbl3	wolinella s	979	38	39.2	335	2	Q5HIF0_STAAC	Q5hif0	staphylococ
907	38	39.2	227	2	Q8NBP8_HUMAN	Q8nbp8	homo sapien	980	38	39.2	336	1	SMT1_ARATH	Q91m02	arabidopsis

908	38	39.2	228	2	Q9VRA4_DROME	Q9vra4	drosophila
909	38	39.2	228	2	Q8SZE4_DROME	Q8sze4	drosophila
910	38	39.2	229	2	Q8KMU6_ENTFA	Q8kmu6	enterococcu
911	38	39.2	229	2	Q88T17_LACPL	Q88t17	lactobacill
912	38	39.2	230	2	Q7ZKL1_9HIV1	Q7zkl1	human immun
913	38	39.2	241	2	Q21302_CAEEL	Q21302	caenorhabdi
914	38	39.2	241	2	O66269_9SPHN	Q66269	erythrombri
915	38	39.2	245	2	O66278_9SPHN	Q66278	agrobacteri
916	38	39.2	245	2	Q9ZN87_9SPHN	Q9zn87	porphyrobac
917	38	39.2	246	2	O66276_9SPHN	Q66276	porphyrobac
918	38	39.2	246	2	Q9I9S7_PHACC	Q9i9s7	phasianus c
919	38	39.2	247	2	O24659_ORYSA	Q24659	oryza sativ
920	38	39.2	247	2	Q8S1X5_ORYSA	Q8s1x5	oryza sativ
921	38	39.2	250	2	Q42648_BETVU	Q42648	beta vulgar
922	38	39.2	250	2	Q88VQ0_LACPL	Q88vq0	lactobacill
923	38	39.2	250	2	O11771_9PARA	Q11771	measles vir
924	38	39.2	250	2	O11776_9PARA	Q11776	measles vir
925	38	39.2	250	2	O11777_9PARA	Q11777	measles vir
926	38	39.2	254	1	AZLC_BACSU	O07942	bacillus su
927	38	39.2	256	1	VMAT_HRSVA	P03419	human respi
928	38	39.2	256	1	Q77YB3_HRSV	Q77yb3	human respi
929	38	39.2	256	2	Q4KRW7_HRSV	Q4krw7	human respi

RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Brickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000172; EAA66385.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 718 AA; 79121 MW; 9E865DCECC4DE39 CRC64;

Query Match 50.0%; Score 48.5; DB 2; Length 718;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 5 PTL-RQWLKSRHTS 18
|||:|||||:
Db 478 PTLRKWLKPRAHTA 492

RESULT 4
Q4UK00 RICFE PRELIMINARY; PRT; 148 AA.
ID Q4UK00 RICFE PRELIMINARY; PRT; 148 AA.
AC Q4UK00;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Uncharacterized conserved protein.
GN OrderedLocusNames=RF_1288;
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
OX NCBI_TaxID=42862;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=URRWXCal2;
RX PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA Ogata H., Renesto P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA Parinello H., Claverie J.-M., Raoult D.;
RT "The genome sequence of Rickettsia felis identifies the first putative
RT conjugative plasmid in an obligate intracellular parasite.";
RL PLoS Biol. 3:E248-E248(2005).
DR EMBL; CP000053; AAY62139.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 148 AA; 17278 MW; 98592DA61DDEE520 CRC64;

Query Match 49.5%; Score 48; DB 2; Length 148;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWL 11
|||||:|:
Db 70 TIKGFSKQWL 80

RESULT 5
P95613 RHIGA PRELIMINARY; PRT; 326 AA.
ID P95613 RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAMBI;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstroem K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.
DR EMBL; Y08963; CAA70157.1; -; Genomic DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LysR; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 49.5%; Score 48; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSR 14
|||:|:
Db 204 KGPSLEQWLSSQ 215

RESULT 6
Q4P0J6 USTMA PRELIMINARY; PRT; 992 AA.
ID Q4P0J6 USTMA PRELIMINARY; PRT; 992 AA.
AC Q4P0J6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM06367.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,

```
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L., Moru K.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Nguyen C.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Ooawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP01000247; EAK87224.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 992 AA; 106443 MW; 2F60D4BA01A2011F CRC64;

Query Match 49.5%; Score 48; DB 2; Length 992;
Best Local Similarity 57.1%; Pred. No. 94;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
   |:|:|:|:|:|:|
Db 561 TVGGPALRMWRKAR 574

RESULT 7
Y745_HELPJ STANDARD; PRT; 327 AA.
AC Q9ZL98;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical RNA pseudouridine synthase JHP0682 (EC 5.4.99.-) (RNA-
DE uridine isomerase) (RNA pseudouridylylate synthase).
GN OrderedLocusNames=JHP0682;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: RNA uridine = RNA pseudouridine.
CC -!- SIMILARITY: Belongs to the pseudouridine synthase rluA family.
CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AE001500; AAD06270.1; -; Genomic_DNA.
DR PIR; B71900; B71900.
DR InterPro; IPR006225; Pseud_synth_RluD.
DR InterPro; IPR006145; PseudoU_synth.
DR InterPro; IPR006224; Rlu_synth.
DR InterPro; IPR002942; S4_RNA_bd.
DR Pfam; PF00849; PseudoU_synth_2; 1.
DR Pfam; PF01479; S4; 1.
DR ProDom; PD001819; PseudoU_synth; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMS; TIGR00005; rluD_subfam; 1.
DR PROSITE; PS01129; PSI_RLU; 1.
DR PROSITE; PS50889; S4; 1.
KW Complete proteome; Hypothetical protein; Isomerase; RNA-binding.
FT DOMAIN 12 79 S4 RNA-binding.
FT ACT SITE 136 136 By similarity.
SQ SEQUENCE 327 AA; 37722 MW; 7EDC7F6840D818BD CRC64;

Query Match 48.5%; Score 47; DB 1; Length 327;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRH 16
   :|:|:|:|:|:|:|
Db 103 SVKEPTLVDWLKSONY 118

RESULT 8
Q750V6 ASHGO
ID Q750V6_ASHGO PRELIMINARY; PRT; 727 AA.
AC Q750V6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGL167Cp.
GN Name=AGL167C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016820; AAS54324.1; -; Genomic_DNA.
DR AGD; AGL167C; -.
KW Complete proteome.
SQ SEQUENCE 727 AA; 82748 MW; 58A66322705F6767 CRC64;

Query Match 48.5%; Score 47; DB 2; Length 727;
Best Local Similarity 56.2%; Pred. No. 97;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRH 18
   :|:|:|:|:|:|:|
Db 92 RGETGRSWRKDRHGS 107

RESULT 9
Q55ZA2 CRYNE
ID Q55ZA2_CRYNE PRELIMINARY; PRT; 1794 AA.
AC Q55ZA2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
```

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBA5860;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AA01000004; EAL23061.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Guanine-nucleotide releasing factor;
KW Hypothetical protein.
SQ SEQUENCE 1794 AA; 196453 MW; 98811FD2F194C994 CRC64;

Query Match 48.5%; Score 47; DB 2; Length 1794;
Best Local Similarity 46.2%; Pred. No. 2.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRH 16
|||:|:|:|:|:
Db 589 GPGIRPQWETRQH 601

RESULT 10
Q5KNL7 CRYNE PRELIMINARY; PRT; 1811 AA.
AC Q5KNL7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ARF guanyl-nucleotide exchange factor, putative.
GN ORFNames=CNA06050;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Krzywinski M.I., Kwon-Chung J.K.,
RA Lengeler K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA Mitchell T.G., Pextea M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA Davis R.W., Fraser C.M., Hyman R.W.;
RT "The genome and transcriptome of Cryptococcus neoformans, a
basidiomycete fungal pathogen of humans.";
RL Science 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung J.K., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Pextea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AE017341; AAW41140.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50190; SEC7; 1.
KW Coiled coil; Complete proteome; Guanine-nucleotide releasing factor.
SQ SEQUENCE 1811 AA; 198207 MW; 3388BA8C9C924019 CRC64;

Query Match 48.5%; Score 47; DB 2; Length 1811;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRH 16
|||:|:|:|:|:
Db 606 GPGIRPQWETRQH 618

RESULT 11
GPD2 MYCPA
ID GPD2 MYCPA STANDARD; PRT; 332 AA.
AC P61744;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Glycerol-3-phosphate dehydrogenase 2 [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-
DE dependent glycerol-3-phosphate dehydrogenase 2).
GN Name=gpsA2; OrderedLocNames=MAP4061c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL -!- CATALYTIC ACTIVITY: sn-glycerol 3-phosphate + NAD(P) (+) =
CC glycerone phosphate + NAD(P)H.
CC -!- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
CC formation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the NAD-dependent glycerol-3-phosphate

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; BA000030; BAC67965.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Complete proteome.
SQ SEQUENCE 402 AA; 43379 MW; 71EC0A91143451BE CRC64;

Query Match 47.4%; Score 46; DB 2; Length 402;
Best Local Similarity 71.4%; Pred. No. 73;
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 2 IKG--PTLRQWLKS 13
Db ||||| |||||
234 IKGNQPTLHQWLKA 247

RESULT 15
Q4KS46_9VIRU
ID Q4KS46_9VIRU PRELIMINARY; PRT; 815 AA.
AC Q4KS46;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tyrosine kinase.
OS Orange-spotted grouper iridovirus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=322017;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhou S.Y., Lv L., Chen C., Weng S.P., Chan S.M., He J.G.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY894343; AAX82420.1; -; Genomic_DNA.
KW Kinase.
SQ SEQUENCE 815 AA; 93153 MW; 120C2FB59703F8C4 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 815;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
Db ::||| |||
500 VQGPTLAQWICSTAF 515

RESULT 16
Q8QUJ6_9VIRU
ID Q8QUJ6_9VIRU PRELIMINARY; PRT; 941 AA.
AC Q8QUJ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF114L.
OS Infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Megalocytivirus.
OX NCBI_TaxID=180170;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus";
RL Virology 291:126-139(2001).
DR EMBL; AF371960; AAL98838.1; -; Genomic_DNA.
SQ SEQUENCE 941 AA; 106703 MW; EB663998C7F6CE83 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 941;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
Db ::||| |||
581 VQGPTLAQWICSTAF 596

RESULT 17
Q980N7_SULSO
ID Q980N7_SULSO PRELIMINARY; PRT; 186 AA.
AC Q980N7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Adenylate cyclase, cyaB-type, putative (CyaB) (EC 4.6.1.1).
GN Name=cyaB; OrderedLocustNames=SSO0253;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006661; AAK40592.1; -; Genomic_DNA.
DR PIR; A90167; A90167.
DR GO; GO:0004016; F:adenylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0006171; P:CAMP biosynthesis; IEA.
DR InterPro; IPR008172; Adenylate_cyc.
DR InterPro; IPR008173; CyaB.
DR Pfam; PF01928; CYTH; 1.
DR PIRSF; PIRSF005720; Ad_cyc_CyaB; 1.
DR ProDom; PD009560; CyaB; 1.
DR TIGRFAMs; TIGR00318; cyaB; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 186 AA; 21820 MW; 1B47C630B438C868 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 186;
Best Local Similarity 55.6%; Pred. No. 45;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSREHTS 18
Db ||||| |||
65 TYKGPKLHSSLKAREEIS 82

RESULT 18
Q8XR40_RALSO
ID Q8XR40_RALSO PRELIMINARY; PRT; 243 AA.
AC Q8XR40;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RSP1028.
GN OrderedLocusNames=RSpl028; ORFNames=RS02365;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646082; CAD18179.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 243 AA; 27220 MW; 2E941BE4DAF5F832 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 243;
Best Local Similarity 61.5%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLRQWLKSRHT 17
Db 149 PGLRNWLSRRQT 161

RESULT 19
Q8T462_DROME
ID Q8T462_DROME PRELIMINARY; PRT; 286 AA.
AC Q8T462;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AT14183p.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089335; AAL90073.1; -; mRNA.
DR Ensembl; CG12717; Drosophila melanogaster.
DR FlyBase; FBgn0063731; BCDNA:AT14183.
SQ SEQUENCE 286 AA; 30787 MW; 99374B2615D88594 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 286;
Best Local Similarity 42.9%; Pred. No. 73;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLRQWLKSRHTS 18
Db 132 PAISQWISRNHAA 145

RESULT 20
Q742B3_MYCPA
ID Q742B3_MYCPA PRELIMINARY; PRT; 302 AA.
AC Q742B3;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE GalU.
GN Name=galU; OrderedLocusNames=MAP0924;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Anonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017230; AAS03241.1; -; Genomic_DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 302;
Best Local Similarity 72.7%; Pred. No. 78;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSR 14
Db 286 GPDLRQWLVAR 296

RESULT 21
Q7VI26_HELHP
ID Q7VI26_HELHP PRELIMINARY; PRT; 351 AA.
AC Q7VI26;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ribosomal pseudouridine synthase.
GN OrderedLocusNames=HH0783;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
CC -!- CATALYTIC ACTIVITY: rRNA uridine = rRNA pseudouridine.
DR EMBL; AE017146; AAP77380.1; -; Genomic_DNA.
DR HSSP; P33643; IQYU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0009982; F:pseudouridine synthase activity; IEA.
DR GO; GO:0004730; F:pseudouridylylate synthase activity; IEA.
DR GO; GO:0003723; F:rRNA binding; IEA.
DR GO; GO:0006364; P:rRNA processing; IEA.
DR InterPro; IPR006145; PseudoU_synth.
DR InterPro; IPR006225; Pseud_synth_Rlud.
DR InterPro; IPR002942; S4.
DR Pfam; PF00849; PseudoU_synth_2; 1.
DR Pfam; PF01479; S4; 1.
DR ProDom; PD001819; PseudoU_synth; 1.
DR SMART; SM00363; S4; 1.

DR TIGRFAMS; TIGR00005; rluD_subfam; 1.
DR PROSITE; PS50889; S4; 1.
KW Complete proteome; Isomerase; RNA-binding; rRNA processing.
SQ SEQUENCE 351 AA; 40281 MW; 9CSB9C3E7733F6DE CRC64;

Query Match 46.4%; Score 45; DB 2; Length 351;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRH 16
: : | | | | | | | |
Db 114 SVKEPTLVDWLKLN 129

RESULT 22

Q6N3E8 RHOPA
ID Q6N3E8 RHOPA PRELIMINARY; PRT; 789 AA.
AC Q6N3E8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Possible RND Superfamily transporter.
GN OrderedLocusNames=RPA3746;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris."
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572605; CAE29187.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000731; SSD_5TM.
DR InterPro; IPR002035; VWF_A.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 789 AA; 85214 MW; 31C1C9E9CAC530CC CRC64;

Query Match 46.4%; Score 45; DB 2; Length 789;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 TLRQWLKSRH 18
| | | | | : | |
Db 523 TLRRLSEKAHTT 535

RESULT 23

Q4QH74 LEIMA
ID Q4QH74 LEIMA PRELIMINARY; PRT; 1126 AA.
AC Q4QH74;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Protein kinase, putative (EC 2.7.1.-).
GN ORFNames=LmjF11.0060;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005250; CAJ02343.1; -; Genomic_DNA.
KW Kinase; Transferase.
SQ SEQUENCE 1126 AA; 124574 MW; 9F72FE6F45D18134 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 1126;
Best Local Similarity 44.4%; Pred. No. 3.4e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRH 18
| : | | | | : | | |
Db 1080 TLAEPTLARWIKANA 1097

RESULT 24

Q5TV94 ANOGA
ID Q5TV94 ANOGA PRELIMINARY; PRT; 1361 AA.
AC Q5TV94;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE ENSANGP0000029021 (Fragment).
GN ORFNames=ENSANGG00000004323;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100844; EAL41356.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR007735; Pecanex_C.
DR Pfam; PF05041; Pecanex_C; 1.
FT NON_TER 1
FT NON_TER 1361
SQ SEQUENCE 1361 AA; 153017 MW; 4578712AEB5107AB CRC64;

Query Match 46.4%; Score 45; DB 2; Length 1361;
Best Local Similarity 37.5%; Pred. No. 4.1e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

QY 1 TIKGPTLRQWLKSR-----EHTS 18
| : | | | | | | | |
Db 726 TLRSPKLSWLSQAIBEALEHTT 749

RESULT 25

Q7PS78 ANOGA
ID Q7PS78 ANOGA PRELIMINARY; PRT; 1431 AA.
AC Q7PS78;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000005643 (Fragment).
GN ORFNames=ENSANGG00000004323;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.

Db	76	VSDPFLREWIKSPEN 90	Query Match Best Local Similarity Matches	45.4%; 72.7%; 8;	Score 44; Pred. No. 1.1e+02; Conservative 1;	DB 2; Length 302; Mismatches 2; Indels 0;	Gaps 0;
RESULT 29							
Q5YPT6	NOCPA						
ID	Q5YPT6	NOCPA PRELIMINARY;	PRT;	298	AA.		
AC	Q5YPT6;						
DT	25-OCT-2004	(TrEMBLrel. 28, Created)					
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)					
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)					
DE	Putative UTP-glucose-1-phosphate uridylyltransferase.						
GN	Name=gauU; OrderedLocusNames=nfa49530;						
OS	Nocardia farcinica.						
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;						
OC	Corynebacterineae; Nocardiaceae; Nocardia.						
OX	NCBI_TaxID=37329;						
RN	[1]						
RP	NUCLEOTIDE SEQUENCE.						
RC	STRAIN=IFM 10152;						
RX	PubMed=15466710; DOI=10.1073/pnas.0406410101;						
RA	Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,						
RA	Shiba T., Hattori M.;						
RT	"The complete genomic sequence of Nocardia farcinica IFM 10152.";						
RL	Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).						
DR	EMBL; AP006618; BAD59805.1; -; Genomic_DNA.						
DR	GO; GO:0016301; F:kinase activity; IEA.						
DR	GO; GO:0016779; F:nucleotidyltransferase activity; IEA.						
DR	GO; GO:0016740; F:transferase activity; IEA.						
DR	GO; GO:0009058; P:biosynthesis; IEA.						
DR	InterPro; IPR005835; NTP_transferase.						
DR	Pfam; PF00483; NTP_transferase; 1.						
KW	Complete proteome; Nucleotidyltransferase; Transferase.						
SQ	SEQUENCE 298 AA; 31931 MW; F29B1A16B26ABA0D CRC64;						
Query Match							
Best Local Similarity							
Matches							
45.4%; Score 44; DB 2; Length 298;							
63.6%; Pred. No. 1.1e+02;							
7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;							
QY	4	GPTLRQWLKSR 14					
Db	275	GPALREWLQHR 285					
RESULT 30							
Q4K6B2	PSEF5						
ID	Q4K6B2	PSEF5 PRELIMINARY;	PRT;	302	AA.		
AC	Q4K6B2;						
DT	13-SEP-2005	(TrEMBLrel. 31, Created)					
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)					
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)					
DE	Hypothetical protein.						
GN	ORFNames=PFL_5142;						
OS	Pseudomonas fluorescens (strain Pf-5).						
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;						
OC	Pseudomonadaceae; Pseudomonas.						
OX	NCBI_TaxID=220664;						
RN	[1]						
RP	NUCLEOTIDE SEQUENCE.						
RC	STRAIN=Pf-5;						
RX	PubMed=15980861; DOI=10.1038/nbt1110;						
RA	Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,						
RA	Mavrodi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,						
RA	Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,						
RA	Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,						
RA	Khouri H.M., Pierson E., Pierson L. III, Thomashow L., Loper J.;						
RT	"Complete genome sequence of the plant commensal Pseudomonas						
RT	fluorescens Pf-5.";						
RL	Nat. Biotechnol. 23:873-878(2005).						
DR	EMBL; CP000076; AAY94364.1; -; Genomic_DNA.						
KW	Hypothetical protein.						
SQ	SEQUENCE 302 AA; 32986 MW; 27F92BC7A9840545 CRC64;						
Query Match							
Best Local Similarity							
Matches							
45.4%; Score 44; DB 2; Length 313;							
61.5%; Pred. No. 1.2e+02;							
8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;							
QY	3	KGPTLRQWLKSR 15					
Db	184	EGPKLRQWPLSKE 196					
RESULT 32							
Q4NFI4	9MICC						
ID	Q4NFI4	9MICC PRELIMINARY;	PRT;	332	AA.		
AC	Q4NFI4;						
DT	13-SEP-2005	(TrEMBLrel. 31, Created)					
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)					
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)					
DE	DTDP-glucose 4,6-dehydratase.						
GN	ORFNames=ArthDRAFT_1942;						
OS	Arthrobacter sp. FB24.						
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;						
OC	Micrococccineae; Micrococcaceae; Arthrobacter.						
OX	NCBI_TaxID=290399;						
RN	[1]						

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of *Arthrobacter* sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of *Arthrobacter* sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -I- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHG01000006; EAL96310.1; -; Genomic_DNA.
DR InterPro; IPR005888; dTDP_gluc_dehyt.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01181; dTDP_gluc_dehyt; 1.
KW NAD.
SQ SEQUENCE 332 AA; 37448 MW; DEC70ABF25BBDE94 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 332;
Best Local Similarity 37.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRHTS 18
|||:|:|:|:
Db 207 KGENVRDWHANDHSS 222

RESULT 33
Q8G7M6 BIFLO
ID Q8G7M6_BIFLO PRELIMINARY; PRT; 347 AA.
AC Q8G7M6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DTDP-glucose 4,6-dehydratase enzyme involved in rhamnose
DE biosynthesis.
GN Name=rmlB1; OrderedLocusNames=BL0229;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of *Bifidobacterium* longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014295; AAN24075.1; -; Genomic_DNA.
DR HSSP; P95780; 1KEP.
DR SMR; Q8G7M6; 17-342.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008460; F:DTDP-glucose 4,6-dehydratase activity; IEA.
DR GO; GO:0051287; F:NAD binding; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR005888; dTDP_gluc_dehyt.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01181; dTDP_gluc_dehyt; 1.
KW Complete proteome; NAD.
SQ SEQUENCE 347 AA; 39388 MW; 34852801FD1334FD CRC64;

Query Match 45.4%; Score 44; DB 2; Length 347;

Best Local Similarity 37.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRHTS 18
|||:|:|:|:
Db 222 KGENVRDWHHTEDHSS 237

RESULT 34
Q53MP5 ORYSA
ID Q53MP5_ORYSA PRELIMINARY; PRT; 371 AA.
AC Q53MP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LOC_Os11g19140;
OS *Oryza sativa* (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; *Oryza*.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Hamilton J.,
RA Jones K., Tallon L., Feldblyum T., Taitrin T., Bera J., Kim M.,
RA Jin S., Fadrosch D., Vuong H., Overton II L., Reardon M., Weaver B.,
RA Johri S., Lewis M., Utterback T., Van Aken S., Wortman J., Haas B.,
RA Koo H., Zismann V., Hsiao J., Iobst S., de Vazeilles A., White O.,
RA Salzberg S., Fraser C.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -I- PATHWAY: Lignin biosynthesis.
DR EMBL; AC135599; AAX96798.1; -; Genomic DNA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR InterPro; IPR000051; SAM_bd.
KW Hypothetical protein; Lignin biosynthesis; Methyltransferase;
KW Transferase.
SQ SEQUENCE 371 AA; 41039 MW; 78C040F7BF64961F CRC64;

Query Match 45.4%; Score 44; DB 2; Length 371;
Best Local Similarity 69.2%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRH 16
|:|:|:|:
Db 131 GETLRESLKRRH 143

RESULT 35
Q82PX5 STRAW
ID Q82PX5_STRAW PRELIMINARY; PRT; 377 AA.
AC Q82PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV747;
OS *Streptomyces avermitilis*.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism *Streptomyces avermitilis*.";

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MFS membrane transporter (Fragment).
GN Name=smt;
OS Gibberella fujikuroi (Bakanae and foot rot disease fungus) (Fusarium moniliforme).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella;
OC Gibberella fujikuroi complex.
OX NCBI_TaxID=5127;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=sm567;
RX MEDLINE=21416226; PubMed=11525413; DOI=10.1007/s002940100218;
RA Voss T., Schulte J., Tudzynski B.;
RT "A new MFS transporter gene next to the gibberellin biosynthesis gene cluster of Gibberella fujikuroi is not involved in gibberellin secretion.";
RL Curr. Genet. 39:377-383 (2001).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AJ272424; CAB75959.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_tr; 1.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRANSPORT.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Transporter.
FT NON TER 1
SQ SEQUENCE 648 AA; 72250 MW; 4C90EEE49E25AF9C CRC64;

Query Match 45.4%; Score 44; DB 2; Length 648;
Best Local Similarity 77.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 QWLKSRHT 17
Db 122 QWLKSQKHT 130

RESULT 39
Q4WKMO ASPFU PRELIMINARY; PRT; 764 AA.
AC Q4WKMO;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE DNA mismatch repair protein Msh4, putative.
GN ORFNames=Afulg02000;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley., Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J., Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S., Farmer M., Fedorova N., Fedorova N., Feidblyum T.V., Fischer R., Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A., Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B., Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J., Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R., Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C., Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M., Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I., Penalta M.A., Pertea M., Price C., Pritchard B.L., Quail M.A., Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M., Ronning C.M., Rutter S., Salzberg S.L., Sanchez M., Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S., Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J., White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K., Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -I- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
CC -I- SIMILARITY: Belongs to the DNA mismatch repair mutS family.
DR EMBL; AAHF01000007; EAL87912.1; -; Genomic_DNA.
DR InterPro; IPR000432; Muts_C.
DR InterPro; IPR007860; Muts_II.
DR InterPro; IPR007696; Muts_III.
DR InterPro; IPR007861; Muts_IV.
DR Pfam; PF05188; Muts_II; 1.
DR Pfam; PF05192; Muts_III; 1.
DR Pfam; PF05190; Muts_IV; 1.
DR Pfam; PF00488; Muts_V; 1.
DR ProDom; PD001263; Muts_C; 1.
DR SMART; SM00534; MUTSac; 1.
DR SMART; SM00533; MUTSd; 1.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA-binding.
SQ SEQUENCE 764 AA; 84832 MW; 8870792DB18E5514 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 764;
Best Local Similarity 58.3%; Pred. No. 3.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLK 12
Db 707 TLEGDALREWLK 718

RESULT 40
Q6VG40 SIVCZ
ID Q6VG40_SIVCZ PRELIMINARY; PRT; 1017 AA.
AC Q6VG40;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Pol protein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22972551; PubMed=14610175;
RX DOI=10.1128/JVI.77.23.12523-12534.2003;
RA Cournaud V., Abela B., Pourrut X., Mpoudi-Ngole E., Loul S., Delaporte E., Peeters M.;
RT "Identification of a new simian immunodeficiency virus lineage with a vpu gene present among different cercopithecus monkeys (C. mona, C. cephus, and C. nictitans) from Cameroon.";
RT J. Virol. 77:12523-12534 (2003).
RL J. Virol. 77:12523-12534 (2003).
CC -I- FUNCTION: Early post-infection, the reverse transcriptase converts the viral RNA genome into double-stranded viral DNA. The RNase H domain of the reverse transcriptase performs two functions. It degrades the RNA template and specifically removes the RNA primer from the RNA/DNA hybrid. Following nuclear import, the integrase catalyzes the insertion of the linear, double-stranded viral DNA into the host cell chromosome. Endogenous Pol proteins may have kept, lost or modified their original function during evolution (By similarity).
CC -I- DOMAIN: The LPQG and YXDD motifs are catalytically important and conserved among many retroviruses (By similarity).
CC -I- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY340701; AAR02377.1; -; Genomic_DNA.

DR HSSP; P12497; 1B9F.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.

DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1017 AA; 114691 MW; A1CFE26C001E6E35 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1017;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKRSRE 15
||| ||| |||
Db 180 EGPKLQWPLSRE 192

RESULT 41
POL_SIVS4
ID POL_SIVS4 STANDARD; PRT; 1019 AA.
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)].
DE Name=POL;
OS Simian immunodeficiency virus (isolate F236/smH4) (SIV-SM).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11737;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89262053; PubMed=2786147; DOI=10.1038/339389a0;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,

RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC !- FUNCTION: During replicative cycle of retroviruses, the reverse-
transcribed viral DNA is integrated into the host chromosome by
the viral integrase enzyme. RNase H activity is associated with
the reverse transcriptase.
CC !- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonooester.
CC !- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
diphosphate + DNA(n+1).
CC !- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC !- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC !- SIMILARITY: Contains 1 integrase catalytic domain.
CC !- SIMILARITY: Contains 1 integrase-type DNA-binding domain.
CC !- SIMILARITY: Contains 1 integrase-type zinc finger.
CC !- SIMILARITY: Contains 1 peptidase A2 domain.
CC !- SIMILARITY: Contains 1 reverse transcriptase domain.
CC !- SIMILARITY: Contains 1 RNase H domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; X14307; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; P12502; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR HIV; X14307; POL\$SMH4.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW AIDS; Aspartyl protease; DNA integration; DNA recombination;
KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
KW Nuclease; Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferrase; Zinc; Zinc-finger.
FT CHAIN 1 167
FT DOMAIN 88 157 Peptidase A2.
FT DOMAIN 211 401 Reverse transcriptase.
FT DOMAIN 600 723 RNase H.
FT DOMAIN 779 930 Integrase catalytic.
FT ZN_FING 729 770 Integrase-type.
FT DNA_BIND 949 996 Integrase-type.
FT ACT_SITE 93 93 By similarity.
SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match

45.4%; Score 44; DB 1; Length 1019;

Best Local Similarity 61.5%; Pred. No. 4.4e+02; Mismatches 2; Indels 0; Gaps 0; Matches 8; Conservative 3;

Qy 3 KGPTLRQWLKRSRE 15 :||| ||||| |:|

Db 184 EGPQLRQWPLSKE 196

RESULT 42

P89154 SIVCZ

ID P89154 SIVCZ PRELIMINARY; PRT; 1019 AA.

AC P89154; 03, Created)

DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Pol protein (Fragment).

GN Name=pol;

OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).

OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;

OC Lentivirus; Primate lentivirus group.

OX NCBI_TaxID=11723;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SIVsmE543;

RX MEDLINE=97151152; PubMed=8995688;

RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,

RA Elkins W.R., Montefiori D.C.;

RT "A molecularly cloned, pathogenic, neutralization-resistant simian

RT immunodeficiency virus, SIVsmE543-3.";

RL J. Virol. 71:1608-1620(1997).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SIVsmE543;

RA Ourmanov I.K., Dehghani H., Kuwata T., Hirsch V.M.;

RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts

CC the viral RNA genome into double-stranded viral DNA. The RNase H

CC domain of the reverse transcriptase performs two functions. It

CC degrades the RNA template and specifically removes the RNA primer

CC from the RNA/DNA hybrid. Following nuclear import, the integrase

CC catalyzes the insertion of the linear, double-stranded viral DNA

CC into the host cell chromosome. Endogenous Pol proteins may have

CC kept, lost or modified their original function during evolution

CC (By similarity).

CC -1- SUBUNIT: Homodimer (By similarity).

CC -1- DOMAIN: The LPQG and YXDD motifs are catalytically important and

CC conserved among many retroviruses (By similarity).

CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.

DR EMBL; U72748; AAC56559.2; -; Genomic_DNA.

DR SMR; P89154; 69-167, 781-939.

DR InterPro; IPR001037; Integrase_C.

DR InterPro; IPR003308; Integrase_N_Zn_bd.

DR InterPro; IPR001969; Pept_Asp_AS.

DR InterPro; IPR009007; Pept_Aspartc_cat.

DR InterPro; IPR001995; Peptidase_A2_cat.

DR InterPro; IPR002156; RNaseH_fold.

DR InterPro; IPR012337; RNaseH_fold.

DR InterPro; IPR001584; Rve.

DR InterPro; IPR010659; RVT_connect.

DR InterPro; IPR010661; RVT_thumb.

DR InterPro; IPR000477; RVTee.

DR InterPro; IPR005829; Sug_transporter.

DR Pfam; PF00552; Integrase; 1.

DR Pfam; PF02022; Integrase_Zn; 1.

DR Pfam; PF00075; RNaseH; 1.

DR Pfam; PF00665; rve; 1.

DR Pfam; PF00077; RVP; 1.

DR Pfam; PF00078; RVT_1; 1.

DR Pfam; PF06815; RVT_connect; 1.

DR Pfam; PF06817; RVT_thumb; 1.

DR PROSITE; PS50175; ASP_PROT_RETROV; 1.

DR PROSITE; PS00141; ASP_PROTEASE; 1.

DR PROSITE; PS50994; INTEGRASE; 1.

DR PROSITE; PS51027; INTEGRASE_DBD; 1.

DR PROSITE; PS50879; RNase_H; 1.

DR PROSITE; PS50878; RT_POL; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.

DR PROSITE; PS50876; ZF_INTEGRASE; 1.

KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;

KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.

FT NON_TER 1

SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;

Best Local Similarity 61.5%; Pred. No. 4.4e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KGPTLRQWLKRSRE 15 :||| ||||| |:|

Db 184 EGPQLRQWPLSKE 196

RESULT 43

Q7ZBR7 SIVCZ

ID Q7ZBR7 SIVCZ PRELIMINARY; PRT; 1019 AA.

AC Q7ZBR7;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Pol (Fragment).

GN Name=pol;

OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).

OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;

OC Lentivirus; Primate lentivirus group.

OX NCBI_TaxID=11723;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22628501; PubMed=12743298;

RX DOI=10.1128/JVI.77.11.6405-6418.2003;

RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;

RT "Unique pattern of convergent envelope evolution in simian

RT immunodeficiency virus-infected rapid progressor macaques: association

RT with CD4-independent usage of CCR5.";

RL J. Virol. 77:6405-6418(2003).

CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts

CC the viral RNA genome into double-stranded viral DNA. The RNase H

CC domain of the reverse transcriptase performs two functions. It

CC degrades the RNA template and specifically removes the RNA primer

CC from the RNA/DNA hybrid. Following nuclear import, the integrase

CC catalyzes the insertion of the linear, double-stranded viral DNA

CC into the host cell chromosome. Endogenous Pol proteins may have

CC kept, lost or modified their original function during evolution

CC (By similarity).

CC -1- SUBUNIT: Homodimer (By similarity).

CC -1- DOMAIN: The LPQG and YXDD motifs are catalytically important and

CC conserved among many retroviruses (By similarity).

CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.

DR EMBL; AY221514; AAC67307.1; -; Genomic_DNA.

DR HSSP; P04584; 1MU2.

DR SMR; Q7ZBR7; 69-167, 781-939.

DR MEROPS; A02.002; -.

DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0008907; F:integrase activity; IEA.

DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004523; F:ribonuclease H activity; IEA.

DR GO; GO:0003723; F:RNA binding; IEA.

DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0015074; P:DNA integration; IEA.

DR GO; GO:0006313; P:DNA transposition; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.

DR InterPro; IPR001037; Integrase_C.

DR InterPro; IPR003308; Integrase_N_Zn_bd.

```
DR InterPro; IPR001969; Pept Asp AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE DBD; 1.
DR PROSITE; PS50879; RNASE H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115341 MW; A886525DFF1BE26F CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRE 15
Db 184 EGPCLRQWPLSKE 196

RESULT 44
Q7ZBR5_SIVCZ
ID Q7ZBR5_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
```

```
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221515; AA067309.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE DBD; 1.
DR PROSITE; PS50879; RNASE H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRE 15
Db 184 EGPCLRQWPLSKE 196

RESULT 45
POL_HV2D2
ID POL_HV2D2 STANDARD; PRT; 1058 AA.
AC P15833;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.47);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)].
GN Name=POL;
OS Human immunodeficiency virus type 2 (isolate D205,7) (HIV-2).
```


OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11716;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=90081881; PubMed=2594088; DOI=10.1038/342948a0;
RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehn H.,
RA Ruebsamen-Waigmann H.;
RT "A highly divergent HIV-2-related isolate.";
RL Nature 342:948-950(1989).
CC -1- FUNCTION: During replicative cycle of retroviruses, the reverse-
transcribed viral DNA is integrated into the host chromosome by
the viral integrase enzyme. RNase H activity is associated with
the reverse transcriptase.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
diphosphate + DNA(n+1).
CC -1- CATALYTIC ACTIVITY: Endopeptidase for which the P1 residue is
preferably hydrophobic.
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC -1- SIMILARITY: Contains 1 integrase catalytic domain.
CC -1- SIMILARITY: Contains 1 integrase-type DNA-binding domain.
CC -1- SIMILARITY: Contains 1 integrase-type zinc finger.
CC -1- SIMILARITY: Contains 1 peptidase A2 domain.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC -1- SIMILARITY: Contains 1 RNase H domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; X61240; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; S08436; S08436.
DR HSSP; P04584; 1MU2.
DR SMR; P15833; 106-204, 207-758, 817-974.
DR HIV; X16109; POL\$2D205.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW AIDS; Aspartyl protease; DNA integration; DNA recombination;
Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
Nuclease; Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferase; Zinc; Zinc-finger.
FT CHAIN 106 204
FT DOMAIN 125 194 Peptidase A2.

FT DOMAIN 248 437 Reverse transcriptase.
FT DOMAIN 636 759 RNase H.
FT DOMAIN 815 966 Integrase catalytic.
FT ZN_FING 765 806 Integrase-type.
FT DNA_BIND 985 1032 Integrase-type.
FT ACT_SITE 130 130 By similarity.
SQ SEQUENCE 1058 AA; 119964 MW; 914D5433694B57F4 CRC64;

Query Match 45.4%; Score 44; DB 1; Length 1058;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRE 15
||:|||||
Db 222 GPKIRQWPLSRE 233

RESULT 46
Q5BFN8 EMENI
ID Q5BFN8 EMENI PRELIMINARY; PRT; 1387 AA.
AC Q5BFN8;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AN0642.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gherre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamat M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AACD01000008; EAA65185.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005786; C:signal recognition particle (sensu Eukaryota); IEA.
DR GO; GO:0008312; F:7S RNA binding; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006605; F:protein targeting; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Transcription; Transcription regulation; Zinc.
SQ SEQUENCE 1387 AA; 152683 MW; 0F30A6D4E00742BD CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1387;
Best Local Similarity 44.4%; Pred. No. 6.2e+02;

		Matches	8;	Conservative	2;	Mismatches	8;	Indels	0;	Gaps	0;																									
QY	1	TIKGPTTLRQWLKSREHTS 18 : : : : : : : : : :																																		
Db	529	TVANVTTRHWMPSEHPS 546																																		
RESULT 47																																				
Q8D674	VIBVU																																			
ID	Q8D674_VIBVU	PRELIMINARY;	PRT;	1928	AA.																															
AC	Q8D674;																																			
DT	01-MAR-2003	(TrEMBLrel. 23, Created)																																		
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)																																		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)																																		
DE	ATP-dependent exoDNase, alpha subunit.																																			
GN	OrderedLocusNames=VV20663;																																			
OS	Vibrio vulnificus.																																			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;																																			
OC	Vibrionaceae; Vibrio.																																			
OX	NCBI_TaxID=672;																																			
RN	[1]																																			
RP	NUCLEOTIDE SEQUENCE.																																			
RC	STRAIN=CMCP6;																																			
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,																																			
RA	Choy H.E.;																																			
RT	"Complete genome sequence of Vibrio vulnificus CMCP6.";																																			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.																																			
DR	EMBL; AE016810; AAO07605.1; -; Genomic_DNA.																																			
DR	HSSP; P14565; 1P4D.																																			
DR	GO; GO:0005524; F:ATP binding; IEA.																																			
DR	GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.																																			
DR	GO; GO:0000166; F:nucleotide binding; IEA.																																			
DR	InterPro; IPR003593; AAA_ATPase.																																			
DR	SMART; SM00382; AAA; 1.																																			
KW	ATP-binding; Complete proteome; Nucleotide-binding; Transport.																																			
SQ	SEQUENCE 1928 AA; 216323 MW; 01D807F50782E486 CRC64;																																			
Query Match 45.4%; Score 44; DB 2; Length 1928;																																				
Best Local Similarity 61.5%; Pred. No. 8.9e+02;																																				
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;																																				
QY	4	GPTLRQWLKSREH 16 : : : : : : : : : :																																		
Db	686	GMTLRQWDKGKPH 698																																		
RESULT 48																																				
Q6MDL6	PARUW																																			
ID	Q6MDL6_PARUW	PRELIMINARY;	PRT;	561	AA.																															
AC	Q6MDL6;																																			
DT	05-JUL-2004	(TrEMBLrel. 27, Created)																																		
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)																																		
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)																																		
DE	Hypothetical protein.																																			
GN	OrderedLocusNames=pc0609;																																			
OS	Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).																																			
OC	Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.																																			
OX	NCBI_TaxID=264201;																																			
RN	[1]																																			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].																																			
RX	PubMed=15073324;																																			
RA	Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,																																			
RA	Fartmann B., Brandt P., Nyakatura G.J., Droege M., Frishman D.,																																			
RA	Rattel T., Mewes H.-W., Wagner M.;																																			
RT	"Illuminating the evolutionary history of chlamydiae.";																																			
RL	Science 304:728-730(2004).																																			
DR	EMBL; BX908798; CAF23333.1; -; Genomic_DNA.																																			
KW	Complete proteome; Hypothetical protein.																																			
SQ	SEQUENCE 561 AA; 63263 MW; 7219CF0E6CA98EDA CRC64;																																			
Query Match 44.8%; Score 43.5; DB 2; Length 561;																																				
Best Local Similarity 50.0%; Pred. No. 2.7e+02;																																				
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;																																				

QY	3	KGP---TLRQWLKSREHT 17																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
----	---	-----------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 130.769 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-80
Perfect score: 97
Sequence: 1 TIKGPTLRQWLKSRHTS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	97	100.0	18	2	AAW09499	Aaw09499 Thrombopo
2	97	100.0	18	2	AAW09459	Aaw09459 Thrombopo
3	97	100.0	18	2	AAW36650	Aaw36650 Thrombopo
4	97	100.0	18	2	AAW33026	Aaw33026 Thrombopo
5	97	100.0	18	3	AAAB17024	Aab17024 TPO-mimet
6	97	100.0	18	4	AAU25869	Aau25869 Human thr
7	97	100.0	18	4	AAU25823	Aau25823 Human thr
8	97	100.0	18	5	ABBB72910	Abb72910 TPO mimet
9	97	100.0	18	7	ADJJ73062	Adj73062 TPO mimet
10	97	100.0	18	8	ADJJ52697	Adj52697 CH1 delet
11	97	100.0	18	8	ADJJ51658	Adj51658 CH1 delet
12	63	64.9	19	2	AAW09491	Aaw09491 Thrombopo
13	63	64.9	19	2	AAW35418	Aaw35418 Thrombopo
14	63	64.9	19	2	AAW36642	Aaw36642 Thrombopo
15	63	64.9	19	4	AAU25861	Aau25861 Human thr
16	63	64.9	19	4	AAU25998	Aau25998 Human thr
17	62	63.9	18	5	ABBP51693	Abp51693 TPO mimet
18	62	63.9	18	5	ABBP51691	Abp51691 TPO mimet
19	62	63.9	18	8	ADQ16625	Adq16625 TPO mimet
20	62	63.9	18	8	ADQ16629	Adq16629 TPO mimet
21	62	63.9	18	9	ADV44363	Adv44363 Agonist T
22	62	63.9	18	9	ADV44367	Adv44367 Agonist T
23	62	63.9	18	9	AEB12840	Aeb12840 TPO mimet
24	62	63.9	18	9	AEB12836	Aeb12836 TPO mimet

98	57	58.8	15	2	AAW36780	Aaw36780 Thrombopo	171	57	58.8	18	8	ADQ16605	Adq16605 TPO mimet
99	57	58.8	15	2	AAW36776	Aaw36776 Thrombopo	172	57	58.8	18	8	ADQ16609	Adq16609 TPO mimet
100	57	58.8	15	2	AAW66714	Aaw66714 Peptide c	173	57	58.8	18	9	ADV44355	Adv44355 Agonist T
101	57	58.8	15	2	AAW66721	Aaw66721 Peptide c	174	57	58.8	18	9	ADV44345	Adv44345 Agonist T
102	57	58.8	15	2	AAW66712	Aaw66712 Peptide c	175	57	58.8	18	9	ADV44351	Adv44351 Agonist T
103	57	58.8	15	3	AAB20684	Aab20684 Thrombocy	176	57	58.8	18	9	ADV44357	Adv44357 Agonist T
104	57	58.8	15	4	AAU25996	Aau25996 Human thr	177	57	58.8	18	9	ADV44349	Adv44349 Agonist T
105	57	58.8	15	4	AAU26026	Aau26026 Human thr	178	57	58.8	18	9	ADV44343	Adv44343 Agonist T
106	57	58.8	15	4	AAU26020	Aau26020 Human thr	179	57	58.8	18	9	ADV44347	Adv44347 Agonist T
107	57	58.8	15	4	AAU25831	Aau25831 Human thr	180	57	58.8	18	9	ADV44384	Adv44384 Agonist T
108	57	58.8	15	4	AAU26007	Aau26007 Human thr	181	57	58.8	18	9	ADV44353	Adv44353 Agonist T
109	57	58.8	15	5	ABP51670	Abp51670 Thrombopo	182	57	58.8	18	9	ADV44361	Adv44361 Agonist T
110	57	58.8	15	7	ABR62908	Abr62908 Thrombopo	183	57	58.8	18	9	ADV44379	Adv44379 Agonist T
111	57	58.8	15	8	ADM72485	Adm72485 TPO mimet	184	57	58.8	18	9	ADV44344	Adv44344 Agonist T
112	57	58.8	15	8	ADM72479	Adm72479 TPO mimet	185	57	58.8	18	9	ADV44359	Adv44359 Agonist T
113	57	58.8	15	8	ADM72502	Adm72502 TPO mimet	186	57	58.8	18	9	ADV44365	Adv44365 Agonist T
114	57	58.8	15	8	ADM72492	Adm72492 TPO mimet	187	57	58.8	18	9	AEB12826	Aeb12826 TPO mimet
115	57	58.8	15	8	ADM72478	Adm72478 TPO mimet	188	57	58.8	18	9	AEB12820	Aeb12820 TPO mimet
116	57	58.8	15	8	ADM72533	Adm72533 TPO mimet	189	57	58.8	18	9	AEB12857	Aeb12857 Antibody
117	57	58.8	15	8	ADM72490	Adm72490 TPO mimet	190	57	58.8	18	9	AEB12834	Aeb12834 TPO mimet
118	57	58.8	15	8	ADM72486	Adm72486 TPO mimet	191	57	58.8	18	9	AEB12830	Aeb12830 TPO mimet
119	57	58.8	15	8	ADM72491	Adm72491 TPO mimet	192	57	58.8	18	9	AEB12832	Aeb12832 TPO mimet
120	57	58.8	15	8	ADM72522	Adm72522 TPO mimet	193	57	58.8	18	9	AEB12852	Aeb12852 TPO mimet
121	57	58.8	15	8	ADM72523	Adm72523 TPO mimet	194	57	58.8	18	9	AEB12816	Aeb12816 TPO mimet
122	57	58.8	15	8	ADM72493	Adm72493 TPO mimet	195	57	58.8	18	9	AEB12818	Aeb12818 TPO mimet
123	57	58.8	15	8	ADM72482	Adm72482 TPO mimet	196	57	58.8	18	9	AEB12828	Aeb12828 TPO mimet
124	57	58.8	15	8	ADQ16585	Adq16585 TPO mimet	197	57	58.8	18	9	AEB12822	Aeb12822 TPO mimet
125	57	58.8	15	8	ADT92483	Adt92483 Modified	198	57	58.8	18	9	AEB12824	Aeb12824 TPO mimet
126	57	58.8	15	9	ADU70209	Adu70209 Thrombopo	199	57	58.8	18	9	AEB12838	Aeb12838 TPO mimet
127	57	58.8	15	9	ADU75981	Adu75981 Peptide-b	200	57	58.8	19	5	ABB73391	Abb73391 TPO-mimet
128	57	58.8	15	9	ADV44320	Adv44320 Agonist T	201	57	58.8	19	5	ABB73390	Abb73390 TPO-mimet
129	57	58.8	15	9	AEB112793	Aeb112793 TPO mimet	202	57	58.8	20	3	AAB18003	Aab18003 Fc-TMP pe
130	57	58.8	16	2	AAB19534	Aaw19534 Thrombopo	203	57	58.8	20	3	AAB17929	Aab17929 TPO-mimet
131	57	58.8	16	2	AAW33035	Aaw33035 Thrombopo	204	57	58.8	20	5	ABB73403	Abb73403 TPO mimet
132	57	58.8	16	2	AAW36775	Aaw36775 Thrombopo	205	57	58.8	21	7	ADN59687	Adn59687 Thrombopo
133	57	58.8	16	2	AAW36771	Aaw36771 Thrombopo	206	57	58.8	22	7	ADN59819	Adn59819 TMP pepti
134	57	58.8	16	2	AAW66709	Aaw66709 Peptide c	207	57	58.8	22	8	ADQ16714	Adq16714 Immunoglo
135	57	58.8	16	2	AAW66713	Aaw66713 Peptide c	208	57	58.8	22	8	ADQ16713	Adq16713 Immunoglo
136	57	58.8	16	2	AAW66733	Aaw66733 Peptide c	209	57	58.8	22	8	ADQ16709	Adq16709 Immunoglo
137	57	58.8	16	2	AAW66716	Aaw66716 Peptide c	210	57	58.8	22	8	ADQ16706	Adq16706 Immunoglo
138	57	58.8	16	4	AAU26021	Aau26021 Human thr	211	57	58.8	22	8	ADQ16699	Adq16699 TPO mimet
139	57	58.8	16	4	AAU26005	Aau26005 Human thr	212	57	58.8	22	8	ADQ16712	Adq16712 Immunoglo
140	57	58.8	16	4	AAU26043	Aau26043 Human thr	213	57	58.8	22	8	ADQ16707	Adq16707 Immunoglo
141	57	58.8	16	4	AAU25832	Aau25832 Human thr	214	57	58.8	22	8	ADQ16711	Adq16711 Immunoglo
142	57	58.8	16	8	ADM72532	Adm72532 TPO mimet	215	57	58.8	22	8	ADQ16708	Adq16708 Immunoglo
143	57	58.8	16	8	ADM72484	Adm72484 TPO mimet	216	57	58.8	22	8	ADQ16710	Adq16710 Immunoglo
144	57	58.8	18	3	AAB16957	Aab16957 PEGylated	217	57	58.8	22	9	ADV44435	Adv44435 Modified
145	57	58.8	18	3	AAB16956	Aab16956 PEGylated	218	57	58.8	22	9	ADV44449	Adv44449 Anti-teta
146	57	58.8	18	5	ABP51687	Abp51687 TPO mimet	219	57	58.8	22	9	ADV44443	Adv44443 Anti-teta
147	57	58.8	18	5	ABP51689	Abp51689 TPO mimet	220	57	58.8	22	9	ADV44444	Adv44444 Anti-teta
148	57	58.8	18	5	ABP51688	Abp51688 TPO mimet	221	57	58.8	22	9	ADV44448	Adv44448 Anti-teta
149	57	58.8	18	5	ABP51677	Abp51677 TPO mimet	222	57	58.8	22	9	ADV44442	Adv44442 Anti-teta
150	57	58.8	18	5	ABP51686	Abp51686 TPO mimet	223	57	58.8	22	9	ADV44447	Adv44447 Anti-teta
151	57	58.8	18	5	ABP51674	Abp51674 TPO mimet	224	57	58.8	22	9	ADV44445	Adv44445 Anti-teta
152	57	58.8	18	5	ABP51684	Abp51684 TPO mimet	225	57	58.8	22	9	ADV44450	Adv44450 Anti-teta
153	57	58.8	18	5	ABP51683	Abp51683 TPO mimet	226	57	58.8	22	9	ADV44446	Adv44446 Anti-teta
154	57	58.8	18	5	ABP51685	Abp51685 TPO mimet	227	57	58.8	22	9	AEB12918	Aeb12918 TPO mimet
155	57	58.8	18	5	ABP51673	Abp51673 TPO mimet	228	57	58.8	22	9	AEB12917	Aeb12917 TPO mimet
156	57	58.8	18	5	ABP51690	Abp51690 TPO mimet	229	57	58.8	22	9	AEB12919	Aeb12919 TPO mimet
157	57	58.8	18	5	ABP51675	Abp51675 TPO mimet	230	57	58.8	22	9	AEB12986	Aeb12986 Tt antibo
158	57	58.8	18	5	ABP51692	Abp51692 TPO mimet	231	57	58.8	22	9	AEB12920	Aeb12920 TPO mimet
159	57	58.8	18	7	ADN59812	Adn59812 Thrombopo	232	57	58.8	22	9	AEB12916	Aeb12916 TPO mimet
160	57	58.8	18	8	ADQ16611	Adq16611 TPO mimet	233	57	58.8	22	9	AEB12914	Aeb12914 TPO mimet
161	57	58.8	18	8	ADQ16619	Adq16619 TPO mimet	234	57	58.8	22	9	AEB12921	Aeb12921 TPO mimet
162	57	58.8	18	8	ADQ16621	Adq16621 TPO mimet	235	57	58.8	22	9	AEB12907	Aeb12907 Tt antibo
163	57	58.8	18	8	ADQ16641	Adq16641 TPO mimet	236	57	58.8	22	9	AEB12915	Aeb12915 TPO mimet
164	57	58.8	18	8	ADQ16646	Adq16646 TPO mimet	237	57	58.8	28	3	AAB17285	Aab17285 TPO-mimet
165	57	58.8	18	8	ADQ16607	Adq16607 TPO mimet	238	57	58.8	28	5	ABP51682	Abp51682 TPO mimet
166	57	58.8	18	8	ADQ16615	Adq16615 TPO mimet	239	57	58.8	28	7	ADJ73013	Adj73013 TPO mimet
167	57	58.8	18	8	ADQ16627	Adq16627 TPO mimet	240	57	58.8	28	8	ADJ52648	Adj52648 CH1 delet
168	57	58.8	18	8	ADQ16617	Adq16617 TPO mimet	241	57	58.8	28	8	ADJ51609	Adj51609 CH1 delet
169	57	58.8	18	8	ADQ16613	Adq16613 TPO mimet	242	57	58.8	28	8	ADQ16636	Adq16636 Tetanus t
170	57	58.8	18	8	ADQ16623	Adq16623 TPO mimet	243	57	58.8	28	9	ADV44374	Adv44374 Modified

244	57	58.8	28	9	AEI12847	Aeb12847	Antibody	317	57	58.8	269	3	AA96531	Human IgG
245	57	58.8	29	3	AAB16975	Aab16975	TPO-mimet	318	57	58.8	269	3	AAB16960	TMP-TMP-F
246	57	58.8	29	3	AAB16976	Aab16976	TPO-mimet	319	57	58.8	269	5	ABB73413	TMP-TMP-F
247	57	58.8	29	3	AAB17286	Aab17286	TPO-mimet	320	57	58.8	472	5	ABP51695	5G1.1-TPO
248	57	58.8	29	3	AAB16970	Aab16970	TPO-mimet	321	57	58.8	472	8	ADQ16647	Immunoglo
249	57	58.8	29	5	ABB72862	Abb72862	TPO mimet	322	57	58.8	472	9	ADV44385	5G1.1 hea
250	57	58.8	29	5	ABB72861	Abb72861	TPO mimet	323	57	58.8	472	9	AEB12858	Antibody
251	57	58.8	29	5	ABB72856	Abb72856	TPO mimet	324	56	57.7	14	2	AAW36773	Thrombopo
252	57	58.8	29	7	ADJ73011	Adj73011	TPO mimet	325	56	57.7	14	8	ADM72527	TPO mimet
253	57	58.8	29	7	ADJ73006	Adj73006	TPO mimet	326	56	57.7	15	2	AAW66731	Peptide c
254	57	58.8	29	8	ADJ52646	Adj52646	CH1 delet	327	56	57.7	15	4	AAU26038	Human thr
255	57	58.8	29	8	ADJ52641	Adj52641	CH1 delet	328	56	57.7	18	7	ADN59663	Thrombopo
256	57	58.8	29	8	ADJ51602	Adj51602	CH1 delet	329	56	57.7	22	7	ADN59830	TMP pepti
257	57	58.8	29	8	ADJ51607	Adj51607	CH1 delet	330	56	57.7	25	7	ADN59708	Thrombopo
258	57	58.8	30	3	AAB17287	Aab17287	TPO-mimet	331	56	57.7	43	7	ADN59759	Peptide-v
259	57	58.8	31	3	AAB17288	Aab17288	TPO-mimet	332	54	55.7	13	4	AAU26008	Human thr
260	57	58.8	31	3	AAB16974	Aab16974	TPO-mimet	333	54	55.7	13	4	AAU26012	Human thr
261	57	58.8	31	3	AAB16973	Aab16973	TPO-mimet	334	54	55.7	14	3	AAB16969	TPO-mimet
262	57	58.8	31	5	ABB72860	Abb72860	TPO mimet	335	54	55.7	14	3	AAB16968	TPO-mimet
263	57	58.8	31	5	ABB72859	Abb72859	TPO mimet	336	54	55.7	14	5	ABB72854	TPO mimet
264	57	58.8	31	7	ADJ73009	Adj73009	TPO mimet	337	54	55.7	14	5	ABB72855	TPO mimet
265	57	58.8	31	7	ADJ73010	Adj73010	TPO mimet	338	54	55.7	14	7	ADJ73005	TPO mimet
266	57	58.8	31	8	ADJ52644	Adj52644	CH1 delet	339	54	55.7	14	7	ADJ73004	TPO mimet
267	57	58.8	31	8	ADJ52645	Adj52645	CH1 delet	340	54	55.7	14	8	ADJ52639	CH1 delet
268	57	58.8	31	8	ADJ51606	Adj51606	CH1 delet	341	54	55.7	14	8	ADJ52640	CH1 delet
269	57	58.8	31	8	ADJ51605	Adj51605	CH1 delet	342	54	55.7	14	8	ADJ51601	CH1 delet
270	57	58.8	32	3	AA96520	Aay96520	Thrombopo	343	54	55.7	14	8	ADJ51600	CH1 delet
271	57	58.8	32	3	AAB17289	Aab17289	TPO-mimet	344	54	55.7	15	2	AAW66718	Peptide c
272	57	58.8	32	3	AAB17297	Aab17297	TPO-mimet	345	54	55.7	29	3	AAB16971	TPO-mimet
273	57	58.8	33	3	AAB17290	Aab17290	TPO-mimet	346	54	55.7	29	3	ABB72857	TPO mimet
274	57	58.8	34	3	AA96527	Aay96527	Thrombopo	347	54	55.7	29	5	ADJ73007	TPO mimet
275	57	58.8	34	3	AAB17291	Aab17291	TPO-mimet	348	54	55.7	29	8	ADJ52642	CH1 delet
276	57	58.8	35	3	AAB17292	Aab17292	TPO-mimet	349	54	55.7	29	8	ADJ51603	CH1 delet
277	57	58.8	36	3	AA96525	Aay96525	Thrombopo	350	53	54.6	12	2	AAW36787	Thrombopo
278	57	58.8	36	3	AA96523	Aay96523	Thrombopo	351	53	54.6	13	4	AAU26035	Human thr
279	57	58.8	36	3	AA96524	Aay96524	Thrombopo	352	53	54.6	13	7	ADJ73003	TPO mimet
280	57	58.8	36	3	AA96526	Aay96526	Thrombopo	353	53	54.6	13	8	ADJ52638	CH1 delet
281	57	58.8	36	3	AAB17307	Aab17307	TPO-mimet	354	53	54.6	13	8	ADJ51599	CH1 delet
282	57	58.8	36	3	AAB17293	Aab17293	TPO-mimet	355	53	54.6	13	8	ADM72524	TPO mimet
283	57	58.8	36	3	AAB16963	Aab16963	TPO-mimet	356	53	54.6	14	2	AAW36788	Thrombopo
284	57	58.8	36	3	AAB17301	Aab17301	TPO-mimet	357	53	54.6	14	4	AAU26013	Human thr
285	57	58.8	36	3	AAB17306	Aab17306	TPO-mimet	358	53	54.6	14	4	AAU26010	Human thr
286	57	58.8	36	3	ABB72403	Abb72403	TPO-mimet	359	53	54.6	15	2	AAW36784	Thrombopo
287	57	58.8	36	5	ABB72403	Abb72403	TPO-mimet	360	53	54.6	15	4	AAU26011	Human thr
288	57	58.8	37	3	AAB17294	Aab17294	TPO-mimet	361	53	54.6	19	2	AAW09457	Thrombopo
289	57	58.8	38	3	AAB17295	Aab17295	TPO-mimet	362	53	54.6	19	2	AAW09492	Thrombopo
290	57	58.8	39	3	AAB17304	Aab17304	TPO-mimet	363	53	54.6	19	2	AAW36651	Thrombopo
291	57	58.8	39	3	AAB17305	Aab17305	TPO-mimet	364	53	54.6	19	2	AAW33024	Thrombopo
292	57	58.8	40	3	AAB17302	Aab17302	TPO-mimet	365	53	54.6	19	2	AAW36643	Thrombopo
293	57	58.8	41	3	AA96528	Aay96528	Thrombopo	366	53	54.6	19	3	AAB17021	TPO-mimet
294	57	58.8	41	5	ABB73389	Abb73389	TPO-mimet	367	53	54.6	19	4	AAU25862	Human thr
295	57	58.8	41	5	ABB73388	Abb73388	TPO-mimet	368	53	54.6	19	4	AAU25870	Human thr
296	57	58.8	42	3	AA96530	Aay96530	Thrombopo	369	53	54.6	19	4	AAU25821	Human thr
297	57	58.8	42	3	AAB17296	Aab17296	TPO-mimet	370	53	54.6	19	5	ABB72907	TPO mimet
298	57	58.8	42	3	AAB17308	Aab17308	Synthetic	371	53	54.6	19	7	ADJ73059	TPO mimet
299	57	58.8	42	3	AAB17282	Aab17282	TPO-mimet	372	53	54.6	19	8	ADJ52694	CH1 delet
300	57	58.8	42	3	AAB17281	Aab17281	TPO-mimet	373	53	54.6	19	8	ADJ51655	CH1 delet
301	57	58.8	42	5	ABB73404	Abb73404	TMP-Tmp g	374	52	53.6	12	2	AAW36781	Thrombopo
302	57	58.8	60	3	AAB17311	Aab17311	Synthetic	375	52	53.6	13	4	AAU26018	Human thr
303	57	58.8	60	3	ABB73405	Abb73405	TMP-TMP g	376	52	53.6	13	8	ADM72525	TPO mimet
304	57	58.8	122	9	ADV44474	Adv44474	Anti-teta	377	52	53.6	13	8	ADM72488	TPO mimet
305	57	58.8	122	9	AEI12946	Aeb12946	Antibody	378	52	53.6	13	8	ADM72489	Thrombopo
306	57	58.8	129	6	ABG71751	Abg71751	Antibody	379	52	53.6	14	2	AAW66715	Peptide c
307	57	58.8	131	6	ABG71753	Abg71753	Antibody	380	52	53.6	14	2	AAW66730	Peptide c
308	57	58.8	133	6	ABG71752	Abg71752	Antibody	381	52	53.6	14	2	AAW66730	Peptide c
309	57	58.8	135	6	ABG71749	Abg71749	Antibody	382	52	53.6	14	4	AAU26009	Human thr
310	57	58.8	143	6	ABG71750	Abg71750	Antibody	383	52	53.6	14	4	AAU26019	Human thr
311	57	58.8	247	3	AAI16958	Aab16958	FC-TMP pr	384	52	53.6	14	4	AAU26036	Human thr
312	57	58.8	247	3	AAB16961	Aab16961	TMP-Fc pr	385	52	53.6	14	8	ADM72495	TPO mimet
313	57	58.8	247	5	ABB73411	Abb73411	FC-TPO m1	386	52	53.6	14	8	ADM72497	TPO mimet
314	57	58.8	247	5	ABB73414	Abb73414	TMP-Fc am	387	52	53.6	15	2	AAW66717	Peptide c
315	57	58.8	268	3	AAI16959	Aab16959	FC-TMP-TM	388	52	53.6	15	4	AAU26022	Human thr
	57	58.8	268	5	ABB73412	Abb73412	FC-TMP-TM	389	52	53.6	15	4	AAU26023	Human thr

390	52	53.6	15	8	ADM72496	Adm72496	TPO mimet	463	47	48.5	8	5	ABP51678	Abp51678	TPO mimet
391	52	53.6	15	8	ADM72494	Adm72494	TPO mimet	464	47	48.5	8	8	ADQ16692	Adq16692	TPO mimet
392	52	53.6	18	7	ADN59681	Adn59681	Thrombopo	465	47	48.5	8	9	ADV44430	Adv44430	Agonist T
393	52	53.6	18	7	ADN59680	Adn59680	Thrombopo	466	47	48.5	8	9	AEB12902	Aeb12902	TPO mimet
394	52	53.6	18	7	ADN59667	Adn59667	Thrombopo	467	47	48.5	14	8	ADM72509	Adm72509	TPO mimet
395	52	53.6	19	2	AAW09494	Aaw09494	Thrombopo	468	47	48.5	15	2	AAW66723	Aaw66723	Peptide c
396	52	53.6	19	2	AAW09461	Aaw09461	Thrombopo	469	47	48.5	15	4	AAU26028	Aau26028	Human thr
397	52	53.6	19	2	AAW33028	Aaw33028	Thrombopo	470	47	48.5	15	8	ADM72508	Adm72508	TPO mimet
398	52	53.6	19	2	AAW36645	Aaw36645	Thrombopo	471	47	48.5	18	5	ABP51679	Abp51679	TPO mimet
399	52	53.6	19	3	AAB17022	Aab17022	TPO-mimet	472	47	48.5	18	7	ADN59672	Adn59672	Thrombopo
400	52	53.6	19	4	AAU25864	Aau25864	Human thr	473	47	48.5	18	7	ADN59658	Adn59658	Thrombopo
401	52	53.6	19	4	AAU25825	Aau25825	Human thr	474	47	48.5	18	7	ADN59675	Adn59675	Thrombopo
402	52	53.6	19	5	ABB72908	Abb72908	TPO mimet	475	47	48.5	18	7	ADN59678	Adn59678	Thrombopo
403	52	53.6	19	7	ADJ73060	Adj73060	TPO mimet	476	47	48.5	18	7	ADN59661	Adn59661	Thrombopo
404	52	53.6	19	8	ADJ52695	Adj52695	CH1 delet	477	47	48.5	18	8	ADQ16693	Adq16693	TPO mimet
405	52	53.6	19	8	ADJ51656	Adj51656	CH1 delet	478	47	48.5	18	9	ADV44431	Adv44431	Anti-teta
406	52	53.6	22	7	ADN59834	Adn59834	TMP pepti	479	47	48.5	18	9	AEB12903	Aeb12903	Heavy cha
407	52	53.6	25	7	ADN59742	Adn59742	Thrombopo	480	47	48.5	22	7	ADN59828	Adn59828	TMP pepti
408	52	53.6	25	7	ADN59716	Adn59716	Thrombopo	481	47	48.5	22	7	ADN59825	Adn59825	TMP pepti
409	52	53.6	25	7	ADN59740	Adn59740	Thrombopo	482	47	48.5	22	7	ADN59839	Adn59839	TMP pepti
410	52	53.6	42	7	ADN59818	Adn59818	Peptide-	483	47	48.5	25	7	ADN59736	Adn59736	Thrombopo
411	52	53.6	43	7	ADN59752	Adn59752	Peptide-v	484	47	48.5	25	7	ADN59704	Adn59704	Thrombopo
412	51	52.6	15	2	AAW66719	Aaw66719	Peptide c	485	47	48.5	25	7	ADN59730	Adn59730	Thrombopo
413	50	51.5	12	4	AAU26014	Aau26014	Human thr	486	47	48.5	25	7	ADN59698	Adn59698	Thrombopo
414	49	50.5	18	7	ADN59815	Adn59815	Thrombopo	487	47	48.5	25	7	ADN59744	Adn59744	Thrombopo
415	49	50.5	18	7	ADN59653	Adn59653	Thrombopo	488	47	48.5	40	7	ADN59753	Adn59753	Peptide-v
416	49	50.5	18	7	ADN59669	Adn59669	Thrombopo	489	46.5	47.9	341	6	ABU33868	Abu33868	Protein e
417	49	50.5	18	7	ADN59673	Adn59673	Thrombopo	490	46	47.4	10	2	AAW09469	Aaw09469	Thrombopo
418	49	50.5	22	7	ADN59820	Adn59820	TMP pepti	491	46	47.4	10	2	AAW36621	Aaw36621	Thrombopo
419	49	50.5	22	7	ADN59840	Adn59840	TMP pepti	492	46	47.4	10	2	AAW36772	Aaw36772	Thrombopo
420	49	50.5	22	7	ADN59836	Adn59836	TMP pepti	493	46	47.4	10	3	AAB17005	Aab17005	TPO-mimet
421	49	50.5	23	7	ADN59778	Adn59778	Peptide-v	494	46	47.4	10	4	AAU25839	Aau25839	Human thr
422	49	50.5	23	7	ADN59796	Adn59796	Peptide-v	495	46	47.4	10	4	AAU25840	Aau25840	Human thr
423	49	50.5	25	7	ADN59726	Adn59726	Thrombopo	496	46	47.4	10	5	ABB72891	Abb72891	TPO mimet
424	49	50.5	25	7	ADN59689	Adn59689	Thrombopo	497	46	47.4	10	7	ADJ73042	Adj73042	TPO mimet
425	49	50.5	25	7	ADN59720	Adn59720	Thrombopo	498	46	47.4	10	8	ADJ52677	Adj52677	CH1 delet
426	49	50.5	36	7	ADN59766	Adn59766	Peptide-v	499	46	47.4	10	8	ADJ51638	Adj51638	CH1 delet
427	49	50.5	36	7	ADN59762	Adn59762	Peptide-v	500	46	47.4	14	8	ADM72499	Adm72499	TPO mimet
428	49	50.5	36	7	ADN59756	Adn59756	Peptide-v	501	46	47.4	15	4	AAU26024	Aau26024	Human thr
429	49	50.5	41	7	ADN59816	Adn59816	Peptide-	502	46	47.4	15	8	ADM72498	Adm72498	TPO mimet
430	49	50.5	41	7	ADN59772	Adn59772	Peptide-v	503	46	47.4	18	2	AAW09487	Aaw09487	Thrombopo
431	49	50.5	46	7	ADN59790	Adn59790	Peptide-v	504	46	47.4	18	2	AAW36638	Aaw36638	Thrombopo
432	49	50.5	46	7	ADN59784	Adn59784	Peptide-v	505	46	47.4	18	4	AAU25857	Aau25857	Human thr
433	48	49.5	10	3	AAB17006	Aab17006	TPO-mimet	506	46	47.4	18	7	ADN59666	Adn59666	Thrombopo
434	48	49.5	10	5	ABB72892	Abb72892	TPO mimet	507	46	47.4	22	7	ADN59833	Adn59833	TMP pepti
435	48	49.5	10	7	ADJ73043	Adj73043	TPO mimet	508	46	47.4	25	7	ADN59714	Adn59714	Thrombopo
436	48	49.5	10	8	ADJ52678	Adj52678	CH1 delet	509	45	46.4	13	3	AAB17015	Aab17015	TPO-mimet
437	48	49.5	10	8	ADJ51639	Adj51639	CH1 delet	510	45	46.4	13	5	ABB72901	Abb72901	TPO mimet
438	48	49.5	12	3	AAB17309	Aab17309	Synthetic	511	45	46.4	13	7	ADJ73054	Adj73054	TPO mimet
439	48	49.5	13	2	AAW36792	Aaw36792	Thrombopo	512	45	46.4	13	7	ADJ73052	Adj73052	TPO mimet
440	48	49.5	13	4	AAU26015	Aau26015	Human thr	513	45	46.4	13	7	ADJ73056	Adj73056	TPO mimet
441	48	49.5	18	7	ADN59660	Adn59660	Thrombopo	514	45	46.4	13	7	ADJ73053	Adj73053	TPO mimet
442	48	49.5	18	7	ADN59654	Adn59654	Thrombopo	515	45	46.4	13	7	ADJ73055	Adj73055	TPO mimet
443	48	49.5	18	7	ADN59655	Adn59655	Thrombopo	516	45	46.4	13	8	ADJ52687	Adj52687	CH1 delet
444	48	49.5	18	7	ADN59670	Adn59670	Thrombopo	517	45	46.4	13	8	ADJ51648	Adj51648	CH1 delet
445	48	49.5	22	7	ADN59821	Adn59821	TMP pepti	518	45	46.4	14	3	AAB17017	Aab17017	TPO-mimet
446	48	49.5	22	7	ADN59827	Adn59827	TMP pepti	519	45	46.4	14	3	AAB17016	Aab17016	TPO-mimet
447	48	49.5	22	7	ADN59822	Adn59822	TMP pepti	520	45	46.4	14	5	ABB72903	Abb72903	TPO mimet
448	48	49.5	22	7	ADN59837	Adn59837	TMP pepti	521	45	46.4	14	5	ABB72902	Abb72902	TPO mimet
449	48	49.5	23	7	ADN59792	Adn59792	Peptide-v	522	45	46.4	14	8	ADJ52689	Adj52689	CH1 delet
450	48	49.5	23	7	ADN59774	Adn59774	Peptide-v	523	45	46.4	14	8	ADJ52688	Adj52688	CH1 delet
451	48	49.5	25	7	ADN59722	Adn59722	Thrombopo	524	45	46.4	14	8	ADJ51650	Adj51650	CH1 delet
452	48	49.5	25	7	ADN59692	Adn59692	Thrombopo	525	45	46.4	14	8	ADJ51649	Adj51649	CH1 delet
453	48	49.5	25	7	ADN59702	Adn59702	Thrombopo	526	45	46.4	15	3	AAB17018	Aab17018	TPO-mimet
454	48	49.5	25	7	ADN59691	Adn59691	Thrombopo	527	45	46.4	15	5	ABB72904	Abb72904	TPO mimet
455	48	49.5	41	7	ADN59768	Adn59768	Peptide-v	528	45	46.4	15	8	ADJ52691	Adj52691	CH1 delet
456	48	49.5	42	7	ADN59751	Adn59751	Peptide-v	529	45	46.4	15	8	ADJ52690	Adj52690	CH1 delet
457	48	49.5	43	7	ADN59761	Adn59761	Peptide-v	530	45	46.4	15	8	ADJ51652	Adj51652	CH1 delet
458	48	49.5	44	7	ADN59817	Adn59817	Peptide-	531	45	46.4	15	8	ADJ51651	Adj51651	CH1 delet
459	48	49.5	46	7	ADN59780	Adn59780	Peptide-v	532	45	46.4	18	7	ADN59664	Adn59664	Thrombopo
460	48	49.5	46	7	ADN59786	Adn59786	Peptide-v	533	45	46.4	18	7	ADN59665	Adn59665	Thrombopo
461	47	48.5	8	2	AAW33037	Aaw33037	Thrombopo	534	45	46.4	18	7	ADN59671	Adn59671	Thrombopo
462	47	48.5	8	4	AAU25982	Aau25982	Human thr	535	45	46.4	18	7	ADN59668	Adn59668	Thrombopo

536	45	46.4	18	7	ADN59674	Adn59674	Thrombopo	609	44	45.4	30	9	ADY64336	Ady64336	Thrombopo
537	45	46.4	18	7	ADN59659	Adn59659	Thrombopo	610	44	45.4	36	3	AAy96521	Aay96521	Cyclic or
538	45	46.4	18	7	ADN59657	Adn59657	Thrombopo	611	44	45.4	36	3	AAB17298	Aab17298	TPO-mimet
539	45	46.4	22	7	ADN59826	Adn59826	TMP pepti	612	44	45.4	36	3	AAB17299	Aab17299	TPO-mimet
540	45	46.4	22	7	ADN59824	Adn59824	TMP pepti	613	44	45.4	52	4	AAU56008	Aau56008	Propionib
541	45	46.4	22	7	ADN59841	Adn59841	TMP pepti	614	44	45.4	52	6	ABM52527	Abm52527	Propionib
542	45	46.4	22	7	ADN59835	Adn59835	TMP pepti	615	44	45.4	100	5	ABP47801	Abp47801	Sooty man
543	45	46.4	22	7	ADN59831	Adn59831	TMP pepti	616	44	45.4	347	5	ABP65398	Abp65398	Bifidobac
544	45	46.4	22	7	ADN59838	Adn59838	TMP pepti	617	44	45.4	429	5	ABP47758	Abp47758	Protein #
545	45	46.4	22	7	ADN59832	Adn59832	Peptide-v	618	43.5	44.8	57	4	AAm94594	Aam94594	Human rep
546	45	46.4	23	7	ADN59776	Adn59776	Peptide-v	619	43	44.3	13	4	AAU26039	Aau26039	Human thr
547	45	46.4	23	7	ADN59797	Adn59797	Peptide-v	620	43	44.3	13	8	ADM72528	Adm72528	TPO mimet
548	45	46.4	23	7	ADN59775	Adn59775	Peptide-v	621	43	44.3	18	2	AAW09488	Aaw09488	Thrombopo
549	45	46.4	23	7	ADN59779	Adn59779	Peptide-v	622	43	44.3	18	2	AAW36639	Aaw36639	Thrombopo
550	45	46.4	23	7	ADN59794	Adn59794	Peptide-v	623	43	44.3	18	4	AAU25858	Aau25858	Human thr
551	45	46.4	23	7	ADN59795	Adn59795	Peptide-v	624	43	44.3	36	3	AAy96522	Aay96522	Linear th
552	45	46.4	23	7	ADN59793	Adn59793	Peptide-v	625	43	44.3	36	3	AAB17300	Aab17300	TPO-mimet
553	45	46.4	23	7	ADN59777	Adn59777	Peptide-v	626	43	44.3	178	4	AAU40877	Aau40877	Propionib
554	45	46.4	25	7	ADN59696	Adn59696	Thrombopo	627	43	44.3	178	6	ABM37396	Abm37396	Propionib
555	45	46.4	25	7	ADN59710	Adn59710	Thrombopo	628	43	44.3	269	6	ABU36398	Abu36398	Protein e
556	45	46.4	25	7	ADN59724	Adn59724	Thrombopo	629	43	44.3	298	3	AAG16591	Aag16591	Arabidops
557	45	46.4	25	7	ADN59712	Adn59712	Thrombopo	630	43	44.3	316	3	AAG16590	Aag16590	Arabidops
558	45	46.4	25	7	ADN59718	Adn59718	Thrombopo	631	43	44.3	369	6	ABU28974	Abu28974	Protein e
559	45	46.4	25	7	ADN59728	Adn59728	Thrombopo	632	43	44.3	375	4	AAU35335	Aau35335	Enterococ
560	45	46.4	25	7	ADN59755	Adn59755	Peptide-v	633	43	44.3	536	4	AAG16589	Aag16589	Arabidops
561	45	46.4	25	7	ADN59700	Adn59700	Thrombopo	634	43	44.3	536	5	ABB93535	Abb93535	Herbicida
562	45	46.4	35	7	ADN59754	Adn59754	Peptide-v	635	43	44.3	536	8	ADN74161	Adn74161	Thale cre
563	45	46.4	36	7	ADN59763	Adn59763	Peptide-v	636	43	44.3	1150	4	AAE11018	Aae11018	Recombina
564	45	46.4	36	7	ADN59765	Adn59765	Peptide-v	637	43	44.3	1150	5	ABB08110	Abb08110	FIV-Oma3
565	45	46.4	36	7	ADN59764	Adn59764	Peptide-v	638	42	43.3	11	2	AAW35425	Aaw35425	Thrombopo
566	45	46.4	36	7	ADN59767	Adn59767	Peptide-v	639	42	43.3	11	4	AAU26001	Aau26001	Human thr
567	45	46.4	41	7	ADN59769	Adn59769	Peptide-v	640	42	43.3	12	2	AAW35423	Aaw35423	Thrombopo
568	45	46.4	41	7	ADN59770	Adn59770	Peptide-v	641	42	43.3	12	4	AAU26000	Aau26000	Human thr
569	45	46.4	41	7	ADN59771	Adn59771	Peptide-v	642	42	43.3	12	8	ADM72530	Adm72530	TPO mimet
570	45	46.4	41	7	ADN59773	Adn59773	Peptide-v	643	42	43.3	13	2	AAW09467	Aaw09467	Thrombopo
571	45	46.4	46	7	ADN59788	Adn59788	Peptide-v	644	42	43.3	13	2	AAW35399	Aaw35399	Thrombopo
572	45	46.4	46	7	ADN59789	Adn59789	Peptide-v	645	42	43.3	13	2	AAW35404	Aaw35404	Thrombopo
573	45	46.4	46	7	ADN59781	Adn59781	Peptide-v	646	42	43.3	13	2	AAW35417	Aaw35417	Thrombopo
574	45	46.4	46	7	ADN59787	Adn59787	Peptide-v	647	42	43.3	13	2	AAW33033	Aaw33033	Thrombopo
575	45	46.4	46	7	ADN59757	Adn59757	Peptide-v	648	42	43.3	13	2	AAW35413	Aaw35413	Thrombopo
576	45	46.4	46	7	ADN59785	Adn59785	Peptide-v	649	42	43.3	13	2	AAW35406	Aaw35406	Thrombopo
577	45	46.4	46	7	ADN59782	Adn59782	Peptide-v	650	42	43.3	13	2	AAW35405	Aaw35405	Thrombopo
578	45	46.4	46	7	ADN59783	Adn59783	Peptide-v	651	42	43.3	13	2	AAW35422	Aaw35422	Thrombopo
579	45	46.4	46	7	ADN59791	Adn59791	Peptide-v	652	42	43.3	13	2	AAW35397	Aaw35397	Thrombopo
580	45	46.4	75	7	ADN59758	Adn59758	Peptide-v	653	42	43.3	13	2	AAW36783	Aaw36783	Thrombopo
581	45	46.4	302	6	ABU33828	Abu33828	Protein e	654	42	43.3	13	4	AAU26041	Aau26041	Human thr
582	44	45.4	14	8	ADM72507	Adm72507	TPO mimet	655	42	43.3	13	4	AAU25994	Aau25994	Human thr
583	44	45.4	14	8	ADM72505	Adm72505	TPO mimet	656	42	43.3	13	4	AAU25997	Aau25997	Human thr
584	44	45.4	14	8	ADM72475	Adm72475	TPO mimet	657	42	43.3	13	4	AAU25991	Aau25991	Human thr
585	44	45.4	14	8	ADM72501	Adm72501	TPO mimet	658	42	43.3	13	4	AAU25990	Aau25990	Human thr
586	44	45.4	14	9	ADU70205	Adu70205	Thrombopo	659	42	43.3	13	4	AAU25984	Aau25984	Human thr
587	44	45.4	14	9	ADU75978	Adu75978	Thrombopo	660	42	43.3	13	8	ADM72529	Adm72529	TPO mimet
588	44	45.4	14	9	ADY64334	Ady64334	Thrombopo	661	42	43.3	14	2	AAW09466	Aaw09466	Thrombopo
589	44	45.4	15	2	AAW66722	Aaw66722	Peptide c	662	42	43.3	14	2	AAW09462	Aaw09462	Thrombopo
590	44	45.4	15	2	AAW66720	Aaw66720	Peptide c	663	42	43.3	14	2	AAW09465	Aaw09465	Thrombopo
591	44	45.4	15	4	AAU25833	Aau25833	Human thr	664	42	43.3	14	2	AAW09482	Aaw09482	Thrombopo
592	44	45.4	15	4	AAU26027	Aau26027	Human thr	665	42	43.3	14	2	AAW33031	Aaw33031	Thrombopo
593	44	45.4	15	4	AAU26025	Aau26025	Human thr	666	42	43.3	14	2	AAW36633	Aaw36633	Thrombopo
594	44	45.4	15	8	ADM72506	Adm72506	TPO mimet	667	42	43.3	14	2	AAW33029	Aaw33029	Thrombopo
595	44	45.4	15	8	ADM72500	Adm72500	TPO mimet	668	42	43.3	14	2	AAW35396	Aaw35396	Thrombopo
596	44	45.4	15	8	ADM72477	Adm72477	TPO mimet	669	42	43.3	14	2	AAW35398	Aaw35398	Thrombopo
597	44	45.4	15	8	ADM72504	Adm72504	TPO mimet	670	42	43.3	14	2	AAW36633	Aaw36633	Thrombopo
598	44	45.4	15	9	ADY64335	Ady64335	Thrombopo	671	42	43.3	14	2	AAW33029	Aaw33029	Thrombopo
599	44	45.4	16	2	AAW66711	Aaw66711	Peptide c	672	42	43.3	14	2	AAW35396	Aaw35396	Thrombopo
600	44	45.4	16	8	ADM72476	Adm72476	TPO mimet	673	42	43.3	14	2	AAW35401	Aaw35401	Thrombopo
601	44	45.4	28	8	ADJ52647	Adj52647	CH1 delet	674	42	43.3	14	2	AAW35403	Aaw35403	Thrombopo
602	44	45.4	29	3	AAB16972	Aab16972	TPO-mimet	675	42	43.3	14	2	AAW36647	Aaw36647	Thrombopo
603	44	45.4	29	5	ABB72858	Abb72858	TPO mimet	676	42	43.3	14	2	AAW35400	Aaw35400	Thrombopo
604	44	45.4	29	7	ADJ73012	Adj73012	TPO mimet	677	42	43.3	14	2	AAW35402	Aaw35402	Thrombopo
605	44	45.4	29	7	ADJ73008	Adj73008	TPO mimet	678	42	43.3	14	2	AAW33032	Aaw33032	Thrombopo
606	44	45.4	29	8	ADJ52643	Adj52643	CH1 delet	679	42	43.3	14	2	AAW66732	Aaw66732	Peptide c
607	44	45.4	29	8	ADJ51604	Adj51604	CH1 delet	680	42	43.3	14	3	AAB17010	Aab17010	TPO-mimet
608	44	45.4	29	8	ADJ51608	Adj51608	CH1 delet	681	42	43.3	14	3	AAB17014	Aab17014	TPO-mimet

682	42	43.3	14	4	AAU25987	Aau25987 Human thr	755	41	42.3	18	4	AAU25867	Aau25867 Human thr
683	42	43.3	14	4	AAU25826	Aau25826 Human thr	756	41	42.3	18	5	ABB72911	Abb72911 TPO mimet
684	42	43.3	14	4	AAU25993	Aau25993 Human thr	757	41	42.3	18	7	ADJ73063	Adj73063 TPO mimet
685	42	43.3	14	4	AAU25852	Aau25852 Human thr	758	41	42.3	18	7	ADN59679	Adn59679 Thrombopo
686	42	43.3	14	4	AAU25989	Aau25989 Human thr	759	41	42.3	18	7	ADN59662	Adn59662 Thrombopo
687	42	43.3	14	4	AAU25983	Aau25983 Human thr	760	41	42.3	18	8	ADJ52698	Adj52698 CH1 delet
688	42	43.3	14	4	AAU25985	Aau25985 Human thr	761	41	42.3	18	8	ADJ51659	Adj51659 CH1 delet
689	42	43.3	14	4	AAU25995	Aau25995 Human thr	762	41	42.3	22	7	ADN59829	Adn59829 TMP pepti
690	42	43.3	14	4	AAU25992	Aau25992 Human thr	763	41	42.3	25	7	ADN59738	Adn59738 Thrombopo
691	42	43.3	14	4	AAU25986	Aau25986 Human thr	764	41	42.3	25	7	ADN59706	Adn59706 Thrombopo
692	42	43.3	14	4	AAU26040	Aau26040 Human thr	765	41	42.3	40	7	ADN59760	Adn59760 Peptide-v
693	42	43.3	14	4	AAU25866	Aau25866 Human thr	766	41	42.3	70	7	ADC88216	Adc88216 Ribosomal
694	42	43.3	14	4	AAU25988	Aau25988 Human thr	767	41	42.3	99	5	ABP47803	Abp47803 FIV prote
695	42	43.3	14	5	ABB72896	Abb72896 TPO mimet	768	41	42.3	184	7	ADH88208	Adh88208 Enterococ
696	42	43.3	14	5	ABB72900	Abb72900 TPO mimet	769	41	42.3	187	3	AAB39464	Aab39464 Gene 26 h
697	42	43.3	14	7	ADJ73047	Adj73047 TPO mimet	770	41	42.3	229	3	AAG34219	Aag34219 Zea may
698	42	43.3	14	7	ADJ73051	Adj73051 TPO mimet	771	41	42.3	244	5	ABG66710	Abg66710 Human nov
699	42	43.3	14	8	ADJ52682	Adj52682 CH1 delet	772	41	42.3	263	5	ABP40153	Abp40153 Staphyloc
700	42	43.3	14	8	ADJ52686	Adj52686 CH1 delet	773	41	42.3	263	8	ADS07199	Ads07199 Staphyloc
701	42	43.3	14	8	ADJ51643	Adj51643 CH1 delet	774	41	42.3	266	7	ADS92500	Ads92500 B. lichen
702	42	43.3	14	8	ADJ51647	Adj51647 CH1 delet	775	41	42.3	276	6	ABU31149	Abu31149 Protein e
703	42	43.3	18	2	AAW09456	Aaw09456 Thrombopo	776	41	42.3	283	8	ADS42544	Ads42544 Bacterial
704	42	43.3	18	2	AAW33023	Aaw33023 Thrombopo	777	41	42.3	295	3	AAG34218	Aag34218 Zea may
705	42	43.3	18	3	AAB17020	Aab17020 TPO-mimet	778	41	42.3	306	5	ABG91624	Abg91624 Purine/py
706	42	43.3	18	4	AAU25820	Aau25820 Human thr	779	41	42.3	306	6	ABU36556	Abu36556 Protein e
707	42	43.3	18	5	ABB72906	Abb72906 TPO mimet	780	41	42.3	306	6	ABU34427	Abu34427 Protein e
708	42	43.3	18	7	ADJ73058	Adj73058 TPO mimet	781	41	42.3	329	8	ADT57646	Adt57646 Plant pol
709	42	43.3	18	7	ADN59677	Adn59677 Thrombopo	782	41	42.3	370	8	ADY11497	Ady11497 Plant ful
710	42	43.3	18	7	ADN59676	Adn59676 Thrombopo	783	41	42.3	376	8	ABM84818	Abm84818 Human dia
711	42	43.3	18	7	ADN59656	Adn59656 Thrombopo	784	41	42.3	376	8	ABM84819	Abm84819 Human dia
712	42	43.3	18	8	ADJ52693	Adj52693 CH1 delet	785	41	42.3	406	3	AAB54065	Aab54065 Human pan
713	42	43.3	18	8	ADJ51654	Adj51654 CH1 delet	786	41	42.3	419	2	AAW01508	Aaw01508 Human pan
714	42	43.3	19	2	AAW09458	Aaw09458 Thrombopo	787	41	42.3	419	2	AAW01509	Aaw01509 Human pan
715	42	43.3	19	2	AAW33025	Aaw33025 Thrombopo	788	41	42.3	419	2	AAW01504	Aaw01504 Wild-type
716	42	43.3	19	4	AAU25822	Aau25822 Human thr	789	41	42.3	419	2	AAW01510	Aaw01510 Human pan
717	42	43.3	22	7	ADN59823	Adn59823 TMP pepti	790	41	42.3	419	2	AAW97618	Aaw97618 Human car
718	42	43.3	25	4	AAU26042	Aau26042 Human thr	791	41	42.3	419	2	AAZ28915	Aay28915 Human reg
719	42	43.3	25	7	ADN59694	Adn59694 Thrombopo	792	41	42.3	419	5	ADI17250	Adi17250 Human NOV
720	42	43.3	25	7	ADN59734	Adn59734 Thrombopo	793	41	42.3	419	5	ADI17249	Adi17249 Human NOV
721	42	43.3	25	7	ADN59732	Adn59732 Thrombopo	794	41	42.3	419	6	ABR54273	Abr54273 Human NOV
722	42	43.3	25	8	ADM72531	Adm72531 TPO mimet	795	41	42.3	419	6	ABR54276	Abr54276 Human NOV
723	42	43.3	57	4	ABG03094	Abg03094 Novel hum	796	41	42.3	419	7	AAE38604	Aae38604 Human car
724	42	43.3	57	4	ABG03096	Abg03096 Novel hum	797	41	42.3	419	8	ADQ30585	Adq30585 Pancreas
725	42	43.3	81	7	ABO75974	AbO75974 Pseudomon	797	41	42.3	426	6	ABR54275	Abr54275 Human NOV
726	42	43.3	140	3	AAB18086	Aab18086 Pinus rad	799	41	42.3	431	5	ABP47760	Abp47760 Protein #
727	42	43.3	180	8	ADX90944	Adx90944 Plant ful	800	41	42.3	492	7	ABO61826	AbO61826 Klebsiell
728	42	43.3	221	7	ABO83112	AbO83112 Pseudomon	801	41	42.3	545	8	ADY12442	Ady12442 Plant ful
729	42	43.3	261	7	ADH86920	Adh86920 Enterococ	802	41	42.3	548	8	ADY08486	Ady08486 Plant ful
730	42	43.3	318	9	ABM91698	Abm91698 M. xanthu	803	41	42.3	569	6	ABU40498	Abu40498 Protein e
731	42	43.3	320	8	ADX78492	Adx78492 Plant ful	804	41	42.3	572	7	ADF06852	Adf06852 Bacterial
732	42	43.3	328	8	ADX68190	Adx68190 Plant ful	805	41	42.3	573	7	ADD95091	Add95091 Chicken g
733	42	43.3	470	9	ABM92990	Abm92990 M. xanthu	806	41	42.3	573	8	ADU59659	Adu59659 Guanylate
734	42	43.3	555	7	ADF05437	Adf05437 Bacterial	807	41	42.3	721	5	ABP26887	Abp26887 Streptoco
735	42	43.3	619	7	ADH87640	Adh87640 Enterococ	808	41	42.3	721	5	ABU46751	Abu46751 Protein e
736	42	43.3	806	4	AAW51617	Aaw51617 Human PDE	809	41	42.3	721	9	AEB91543	Aeb91543 Microbial
737	42	43.3	1061	2	AAR04192	Aar04192 Pol gene	810	41	42.3	762	4	ABG08032	Abg08032 Novel hum
738	42	43.3	1239	6	ABU18757	Abu18757 Protein e	811	41	42.3	1055	3	AAB12993	Aab12993 HIV-2 rev
739	42	43.3	1241	6	ABU18092	Abu18092 Protein e	812	41	42.3	3437	7	ABB66639	Abb66639 Drosophil
740	42	43.3	4051	7	ADM77465	Adm77465 Rat fibro	813	40.5	42.3	520	7	ABO64867	AbO64867 Klebsiell
741	41.5	42.8	110	4	AAW85655	Aaw85655 Human imm	814	40.5	41.8	569	7	ADC68464	Adc68464 Lolium pe
742	41.5	42.8	168	4	ABB65448	Abb65448 Drosophil	815	40.5	41.8	569	9	AEBO3167	Aeb03167 Fructan b
743	41.5	42.8	252	6	ABU17253	Abu17253 Protein e	816	40.5	41.8	571	5	ABG69061	Abg69061 Amino aci
744	41.5	42.8	262	6	ADA34120	Ada34120 Acinetoba	817	40	41.2	15	9	ADV21762	Adv21762 SIV pol p
745	41.5	42.8	308	4	AAG91969	Aag91969 C glutami	818	40	41.2	18	2	AAW09490	Aaw09490 Thrombopo
746	41	42.3	10	2	AAW09472	Aaw09472 Thrombopo	819	40	41.2	18	2	AAW36641	Aaw36641 Thrombopo
747	41	42.3	10	2	AAW36623	Aaw36623 Thrombopo	820	40	41.2	18	4	AAU25860	Aau25860 Human thr
748	41	42.3	10	4	AAU25842	Aau25842 Human thr	821	40	41.2	19	2	AAW09495	Aaw09495 Thrombopo
749	41	42.3	15	2	AAW66724	Aaw66724 Peptide c	822	40	41.2	19	2	AAW36646	Aaw36646 Thrombopo
750	41	42.3	18	2	AAW09497	Aaw09497 Thrombopo	823	40	41.2	19	3	AAB17023	Aab17023 TPO-mimet
751	41	42.3	18	2	AAW36653	Aaw36653 Thrombopo	824	40	41.2	19	4	AAU25865	Aau25865 Human thr
752	41	42.3	18	2	AAW36648	Aaw36648 Thrombopo	825	40	41.2	19	5	ABB72909	Abb72909 TPO mimet
753	41	42.3	18	3	AAB17025	Aab17025 TPO-mimet	826	40	41.2	19	7	ADJ73061	Adj73061 TPO mimet
754	41	42.3	18	4	AAU25872	Aau25872 Human thr	827	40	41.2	19	8	ADJ52696	Adj52696 CH1 delet

828	40	41.2	19	8	ADJ51657	Adj51657	CH1 delet	901	40	41.2	1060	2	AAW89314	Aaw89314	SIVmac239
829	40	41.2	28	7	AAE38642	Aae38642	Human col	902	40	41.2	1061	8	ABM82874	Abm82874	Human dia
830	40	41.2	87	4	AAU63256	Aau63256	Propionib	903	40	41.2	1078	4	AAM40012	Aam40012	Human pol
831	40	41.2	87	6	ABM59775	Abm59775	Propionib	904	40	41.2	1104	4	ABB67107	Abb67107	Drosophil
832	40	41.2	100	5	ABP47799	Abp47799	SIV prote	905	40	41.2	1118	9	ADV23765	Adv23765	SIV full
833	40	41.2	100	5	ABP47800	Abp47800	HIV-2prot	906	40	41.2	1712	5	AAO17361	Aao17361	Human alp
834	40	41.2	108	4	AAU57517	Aau57517	Propionib	907	40	41.2	1712	8	ADS10473	Ads10473	Human the
835	40	41.2	108	6	ABM54036	Abm54036	Propionib	908	40	41.2	1712	8	ADU06526	Adu06526	Novel bro
836	40	41.2	139	4	AAU57710	Aau57710	Propionib	909	40	41.2	1712	9	ADY59885	Ady59885	Human col
837	40	41.2	139	6	ABM54229	Abm54229	Propionib	910	40	41.2	10625	9	ADZ51727	Adz51727	PKS FscC,
838	40	41.2	183	7	ABO77298	AbO77298	Pseudomoi	911	39.5	40.7	124	7	ADF61805	Adf61805	Bovine D1
839	40	41.2	184	7	ADG42397	Adg42397	SSP domai	912	39.5	40.7	202	5	AAU92979	Aau92979	Arabidops
840	40	41.2	200	8	ADN23470	Adn23470	Bacterial	913	39.5	40.7	202	6	AAE30377	Aae30377	Arabidops
841	40	41.2	201	7	ADG42384	Adg42384	Trypanoso	914	39.5	40.7	202	7	ADD30973	Add30973	Plant yie
842	40	41.2	230	2	AAR60767	Aar60767	HIV-2 ROD	915	39.5	40.7	202	7	ADE37075	Ade37075	Plant yie
843	40	41.2	242	6	ABU34805	Abu34805	Protein e	916	39.5	40.7	202	8	ADI41879	Adi41879	Plant tra
844	40	41.2	242	6	ABU36571	Abu36571	Protein e	917	39.5	40.7	202	8	ADO02529	Ado02529	Thalecres
845	40	41.2	244	8	ADQ67688	Adq67688	Novel hum	918	39.5	40.7	202	9	AEA26403	Aea26403	Stress to
846	40	41.2	261	6	AAO23118	Aao23118	'Human mo	919	39.5	40.7	240	7	ADF61804	Adf61804	Bovine D1
847	40	41.2	274	4	ABG09346	Abg09346	Novel hum	920	39.5	40.7	319	8	ADX97097	Adx97097	Plant ful
848	40	41.2	282	8	ADS42572	Ads42572	Bacterial	921	39.5	40.7	475	4	ABB63526	Abb63526	Drosophil
849	40	41.2	282	8	ADS24597	Ads24597	Bacterial	922	39.5	40.7	643	7	ABO69736	Abo69736	Pseudomon
850	40	41.2	284	9	ABM91232	Abm91232	M. xanthu	923	39	40.2	14	8	ADM72511	Adm72511	TPO mimet
851	40	41.2	306	5	ABG91631	Abg91631	Purine/py	924	39	40.2	14	8	ADM72519	Adm72519	TPO mimet
852	40	41.2	306	6	ABU35738	Abu35738	Protein e	925	39	40.2	15	2	AAW66728	Aaw66728	Peptide c
853	40	41.2	309	3	AAB25283	Aab25283	Eucalyptu	926	39	40.2	15	4	AAU26033	Aau26033	Human thr
854	40	41.2	311	4	AAM41798	Aam41798	Human pol	927	39	40.2	15	4	AAU26029	Aau26029	Human thr
855	40	41.2	316	9	ABM93583	Abm93583	M. xanthu	928	39	40.2	15	8	ADM72510	Adm72510	TPO mimet
856	40	41.2	342	7	ABO83666	Abo83666	Pseudomon	929	39	40.2	15	8	ADM72518	Adm72518	TPO mimet
857	40	41.2	349	7	ADC96606	Adc96606	E. faeciu	930	39	40.2	18	2	AAW09592	Aaw09592	Thrombopo
858	40	41.2	349	8	ADK47468	Adk47468	Streptoco	931	39	40.2	18	2	AAW36743	Aaw36743	Thrombopo
859	40	41.2	352	8	ADR94598	Adr94598	Novel S.	932	39	40.2	18	4	AAU25962	Aau25962	Human thr
860	40	41.2	352	9	AEA59852	Aea59852	Streptoco	933	39	40.2	70	6	AAO16789	Aao16789	Chicken n
861	40	41.2	353	8	ADS42696	Ads42696	Bacterial	934	39	40.2	95	5	AAU09166	Aau09166	Human cyt
862	40	41.2	355	6	ABR58696	Abr58696	Human can	935	39	40.2	102	4	AAU31672	Aau31672	Novel hum
863	40	41.2	355	7	ADF13714	Adf13714	Tumor-Ass	936	39	40.2	106	4	AAE04133	Aae04133	Human gen
864	40	41.2	355	8	ADS00100	Ads00100	Human ocu	937	39	40.2	106	5	ABG64355	Abg64355	Human alb
865	40	41.2	355	8	ADS00098	Ads00098	Human ocu	938	39	40.2	106	8	ADL77622	Adl77622	Albumin f
866	40	41.2	355	9	ADW14773	Adw14773	Tumor-ass	939	39	40.2	109	4	ABG00345	Abg00345	Novel hum
867	40	41.2	361	4	AAU64571	Aau64571	Propionib	940	39	40.2	126	7	ADM03803	Adm03803	Human pro
868	40	41.2	361	6	ABM61090	Abm61090	Propionib	941	39	40.2	184	4	ABG00073	Abg00073	Novel hum
869	40	41.2	361	8	ADL70335	Adl70335	Crenarcha	942	39	40.2	202	6	ABO01277	Abo01277	Growth de
870	40	41.2	368	4	AAU33841	Aau33841	Staphyloc	943	39	40.2	214	5	ABP03213	Abp03213	Human ORF
871	40	41.2	401	6	ABU23114	Abu23114	Protein e	944	39	40.2	227	7	ADF05977	Adf05977	Bacterial
872	40	41.2	428	5	ABP47756	Abp47756	Protein #	945	39	40.2	230	3	AAB56838	Aab56838	Human pro
873	40	41.2	430	5	ABP47756	Abp47756	Protein #	946	39	40.2	235	6	ABM70506	Abm70506	Photorhab
874	40	41.2	476	4	ABG12329	Abg12329	Novel hum	947	39	40.2	238	9	ABM93884	Abm93884	M. xanthu
875	40	41.2	476	4	ABG02325	Abg02325	Novel hum	948	39	40.2	240	5	ABB48752	Abb48752	Listeria
876	40	41.2	476	4	ABG12749	Abg12749	Novel hum	949	39	40.2	269	4	ABG12414	Abg12414	Novel hum
877	40	41.2	498	8	ADI45117	Adi45117	Rice isop	950	39	40.2	273	8	ADX74092	Adx74092	Plant ful
878	40	41.2	549	2	AAW40113	Aaw40113	Human alp	951	39	40.2	290	8	ADU02357	Adu02357	Novel hum
879	40	41.2	603	8	ADS25043	Ads25043	Bacterial	952	39	40.2	336	2	AAW06488	Aaw06488	Maize Ts2
880	40	41.2	653	4	ABE64562	Abb64562	Drosophil	953	39	40.2	336	8	ADT58063	Adt58063	Plant pol
881	40	41.2	896	6	ABU28940	Abu28940	Protein e	954	39	40.2	349	5	ABB53502	Abb53502	Lactococc
882	40	41.2	898	4	ABU64712	Abb64712	Drosophil	955	39	40.2	349	8	ADS29275	Ads29275	Bacterial
883	40	41.2	978	6	ABU16227	Abu16227	Protein e	956	39	40.2	350	2	AAW62621	Aaw62621	Mus muscu
884	40	41.2	981	4	AAU36618	Aau36618	Staphyloc	957	39	40.2	351	8	ADY13095	Ady13095	Plant ful
885	40	41.2	981	6	ABM71093	Abm71093	Staphyloc	958	39	40.2	358	8	ADY10094	Ady10094	Plant ful
886	40	41.2	994	8	ABM82875	Abm82875	Human dia	959	39	40.2	364	6	ABU23456	Abu23456	Protein e
887	40	41.2	1000	9	ADX40089	Adx40089	HIV Pol p	960	39	40.2	397	6	ABR82135	Abr82135	Bacterial
888	40	41.2	1009	2	AAR10275	Aar10275	Simian im	961	39	40.2	398	8	ABO58560	Abo58560	Human gen
889	40	41.2	1009	8	ADN23142	Adn23142	Bacterial	962	39	40.2	492	7	ABO82111	Abo82111	Pseudomon
890	40	41.2	1014	1	AAP80810	Aap80810	Sequence	963	39	40.2	492	9	ADX39448	Adx39448	HIV Gag p
891	40	41.2	1027	1	AAP81773	Aap81773	Sequence	964	39	40.2	496	9	ABM91663	Abm91663	M. xanthu
892	40	41.2	1035	2	AAR04025	Aar04025	Pol gene	965	39	40.2	497	4	AAB66268	Aab66268	Murine TA
893	40	41.2	1036	2	AAR20599	Aar20599	ROD HIV-2	966	39	40.2	562	4	ABB59914	Abb59914	Drosophil
894	40	41.2	1036	3	AAAY51978	Aay51978	HIV-2 ROD	967	39	40.2	567	2	AAAY23855	Aay23855	A Trichom
895	40	41.2	1036	6	ADA07961	Ada07961	HIV-2 pol	968	39	40.2	571	4	AAG70747	Aag70747	S cerevis
896	40	41.2	1036	8	ADL23699	Adl23699	HIV-2 ROD	969	39	40.2	621	7	ABM87339	Abm87339	Rice abio
897	40	41.2	1040	6	ABP81250	Abp81250	Arabidops	970	39	40.2	623	7	ABO65294	Abo65294	Kiebsiell
898	40	41.2	1056	1	AAP81783	Aap81783	Sequence	971	39	40.2	656	8	ADJ71949	Adj71949	Human PMM
899	40	41.2	1056	1	AAP80809	Aap80809	Sequence	972	39	40.2	663	8	ADS42849	Ads42849	Bacterial
900	40	41.2	1060	2	AAR22366	Aar22366	SIVmac239	973	39	40.2	678	3	AAG30788	Aag30788	Arabidops

901	40	41.2	1060	2	AAW89314	Aaw89314	SIVmac239
902	40	41.2	1061	8	ABM82874	Abm82874	Human dia
903	40	41.2	1078	4	AAM40012	Aam40012	Human pol
904	40	41.2	1104	4	ABB67107	Abb67107	Drosophil
905	40	41.2	1118	9	ADV23765	Adv23765	SIV full
906	40	41.2	1712	5	AAO17361	Aao17361	Human alp
907	40	41.2	1712	8	ADS10473	Ads10473	Human the
908	40	41.2	1712	8	ADU06526	Adu06526	Novel bro
909	40	41.2	1712	9	ADY59885	Ady59885	Human col
910	40	41.2	10625	9	ADZ51727	Adz51727	PKS FscC,
911	39.5	40.7	124	7	ADF61805	Adf61805	Bovine D1
912	39.5	40.7	202	5	AAU92979	Aau92979	Arabidops
913	39.5	40.7	202	6	AAE30377	Aae30377	Arabidops
914	39.5	40.7	202	7	ADD30973	Add30973	Plant yie
915	39.5	40.7	202	7	ADE37075	Ade37075	Plant yie
916	39.5	40.7	202	8	ADI41879	Adi41879	Plant tra
917	39.5	40.7	202	8	ADO02529	Ado02529	Thalecres
918	39.5	40.7	202	9	AEA26403	Aea26403	Stress to
919	39.5	40.7	240	7	ADF61804	Adf61804	Bovine D1
920	39.5	40.7	319	8	ADX97097	Adx97097	Plant ful
921	39.5	40.7	475	4	ABB63526	Abb63526	Drosophil
922	39.5	40.7	643	7	ABO69736	Abo69736	Pseudomon
923	39	40.2	14	8	ADM72511	Adm72511	TPO mimet
924	39	40.2	14	8	ADM72519	Adm72519	TPO mimet
925	39	40.2	15	2	AAW66728	Aaw66728	Peptide c
926	39	40.2	15	4	AAU26033	Aau26033	Human thr
927	39	40.2	15	4	AAU26029	Aau26029	Human thr
928	39	40.2	15	8	ADM72510	Adm72510	TPO mimet
929	39	40.2	15	8	ADM72518	Adm72518	TPO mimet
930	39	40.2	18	2	AAW09592	Aaw09592	Thrombopo
931	39	40.2	18	2	AAW36743	Aaw36743	Thrombopo
932	39	40.2	18	4	AAU25962	Aau25962	Human thr
933	39	40.2	70	6	AAO16789	Aao16789	Chicken n
934	39	40.2	95	5	AAU09166	Aau09166	Human cyt
935	39	40.2	102	4	AAU31672	Aau31672	Novel hum
936	39	40.2	106	4	AAE04133	Aae04133	Human gen
937	39	40.2	106	5	ABG64355	Abg64355	Human alb
938	39	40.2	106	8	ADL77622	Adl77622	Albuminf
939	39	40.2	109	4	ABG00345	Abg00345	Novel hum
940	39	40.2	126	7	ADM03803	Adm03803	Human pro
941	39	40.2	184	4	ABG00073	Abg00073	Novel hum
942	39	40.2	202	6	ABO01277	Abo01277	Growth de
943	39	40.2	214	5	ABP03213	Abp03213	Human ORF
944	39	40.2	227	7	ADF05977	Adf05977	Bacterial
945	39	40.2	230	3	AAB56838	Aab56838	Human pro
946	39	40.2	235	6	ABM70506	Abm70506	Photorhab
947	39	40.2	238	9	ABM93884	Abm93884	M. xanthu
948	39	40.2	240	5	ABB48752	Abb48752	Listeria
949	39	40.2	269	4	ABG12414	Abg12414	Novel hum
950	39	40.2	273	8	ADX74092	Adx74092	Plant ful
951	39	40.2	290	8	ADU02357	Adu02357	Novel hum
952	39	40.2	336	2	AAW06488	Aaw06488	Maize Ts2
953	39	40.2	336	8	ADT58063	Adt58063	Plant pol
954	39	40.2	349	5	ABB53502	Abb53502	Lactococc
955	39	40.2	349	8	ADS29275	Ads29275	Bacterial
956	39	40.2	350	2	AAW62621	Aaw62621	Mus muscu
957	39	40.2	351	8	ADY13095	Ady13095	Plant ful
958	39	40.2	358	8	ADY10094	Ady10094	Plant ful
959	39	40.2	364	6	ABU23456	Abu23456	Protein e
960	39	40.2	397	6	ABR82135	Abrr82135	Bacterial
961	39	40.2	398	8	ABO58560	Abo58560	Human gen
962	39	40.2	492	7	ABO82111	Abo82111	Pseudomon
963	39	40.2	492	9	ADX39448	Adx39448	HIV Gag p
964	39	40.2	496	9	ABM91663	Abm91663	M. xanthu
965	39	40.2	497	4	AAB66268	Aab66268	Murine TA
966	39	40.2	562	4	ABB59914	Abb59914	Drosophil
967	39	40.2	567	2	AAZ23855	Aay23855	A Trichom
968	39	40.2	571	4	AAG70747	Aag70747	S cerevis
969	39	40.2	621	7	ABM87339	Abm87339	Rice abio
970	39	40.2	623	7	ABO65294	Abo65294	Klebsiell
971	39	40.2	656	8	ADJ71949	Adj71949	Human PMM
972	39	40.2	663	8	ADS42849	Ads42849	Bacterial
973	39	40.2	678	3	AAG30788	Aag30788	Arabidops

DR WPI; 1997-051883/05.
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 18; Page 89; 106pp; English.
XX
CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding
CC affinity to TR as expressed by an IC50 of no more than about 100 mum. The
CC compound (especially if modified, see features table) can be used for
CC treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18
Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 3
AAW36650
ID AAW36650 standard; peptide; 18 AA.
XX
AC AAW36650;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
AC AAW36650;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 27; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18
Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 4
AAW33026
ID AAW33026 standard; peptide; 18 AA.
XX
AC AAW33026;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 19; Page 89; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microM. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TIKGPTLRQWLKSRHTS 18
Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 7
AAU25823
ID AAU25823 standard; peptide; 18 AA.
XX AC
XX AAU25823;
DT 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #9.
DE
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX Homo sapiens.
XX US6251864-B1.
PN 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
DR
XX Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
PT
XX Disclosure; Col 67-68; 128pp; English.
PS
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX Sequence 18 AA;
SQ
Query Match 100.0%; Score 97; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSRHTS 18
Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 8
ABB72910
ID ABB72910 standard; peptide; 18 AA.
XX AC
XX ABB72910;
DT 05-APR-2002 (first entry)
XX TPO mimetic peptide SEQ ID NO:80.
DE
XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX Homo sapiens.
OS Synthetic.
XX WO200183525-A2.
PN 08-NOV-2001.
XX
XX 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
PI WPI; 2002-130313/17.
DR
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
XX Claim 39; Page 44; 176pp; English.
PS
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal protease of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX Sequence 18 AA;
SQ

Query Match 100.0%; Score 97; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSREHTS 18
Db 1 TIKGPTLRQWLKSREHTS 18

RESULT 9

ADJ73062
ID ADJ73062 standard; peptide; 18 AA.
XX
AC ADJ73062;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 516.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.

XX WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX DR WPI; 2003-804237/75.

XX PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.

PS Disclosure; SEQ ID NO 516; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSREHTS 18
Db 1 TIKGPTLRQWLKSREHTS 18

RESULT 10
ADJ52697
ID ADJ52697 standard; peptide; 18 AA.
XX
AC ADJ52697;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID516.

XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX OS Unidentified.
OS Synthetic.

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

XX PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX PS Claim 2; SEQ ID NO 516; 129pp; English.

XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSREHTS 18
Db 1 TIKGPTLRQWLKSREHTS 18

RESULT 11

ADJ51658
ID ADJ51658 standard; peptide; 18 AA.
XX
AC ADJ51658;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID516.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 516; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 8; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSRHTS 18
| | | | | | | | | | | | | | | | | |
Db 1 TIKGPTLRQWLKSRHTS 18

RESULT 12
AAW09491
ID AAW09491 standard; protein; 19 AA.
XX
AC AAW09491;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.

Thrombopoietin receptor-binding/activating peptide(s) and peptide
mimetic(s) - useful in treatment of haematological disorders, esp.
thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 26; 106pp; English.

The present sequence is a peptide which binds to thrombopoietin (TPO)
receptor (TR). The compound can be used for treating patients suffering
from haematological disorders and thrombocytopenia resulting from
chemotherapy, radiation therapy or bone marrow transfusions. The peptide
may also be used to maintain the proliferation and growth of TPO-
dependent cell lines and for use in biological research, for detecting
TPO receptors on living cells

Sequence 19 AA;

Query Match 64.9%; Score 63; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRHTS 18
| | | | | | | | | | | | | | | | | |
Db 5 GPTLRQWLAAARNHLS 19

RESULT 13
AAW35418
ID AAW35418 standard; peptide; 19 AA.
XX
AC AAW35418;

XX 11-MAR-1998 (first entry)
DT
XX
DE
XX
XX
KW Thrombopoietin receptor binding peptide.
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Cross-links 3
FT Modified-site /note= "linked via disulfide bond to Cys3 of identical peptide"
FT 19
FT /note= "NH2-Ser"
XX
PN WO9640750-A1.
XX
XX 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Example 9; Page 73; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 19 AA;

Query Match 64.9%; Score 63; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRHTS 18
| | | | | : | | |
Db 5 GPTLRQWLAARNHLS 19

RESULT 14
AAW36642
ID AAW36642 standard; peptide; 19 AA.
XX
AC AAW36642;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
XX 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 19 AA;

Query Match 64.9%; Score 63; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSRHTS 18
| | | | | : | | |
Db 5 GPTLRQWLAARNHLS 19

RESULT 15
AAU25861
ID AAU25861 standard; peptide; 19 AA.
XX
AC AAU25861;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #47.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 20; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent hematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 19 AA;
Query Match 64.9%; Score 63; DB 4; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSRHTS 18
| | | | | : | | |
Db 5 GPTLRQWLAARNHLS 19
RESULT 16
AAU25998
ID AAU25998 standard; peptide; 19 AA.
XX
AC AAU25998;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #184.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; hematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 143-144; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent hematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 19 AA;
Query Match 64.9%; Score 63; DB 4; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSRHTS 18
| | | | | : | | |
Db 5 GPTLRQWLAARNHLS 19
RESULT 17
ABP51693
ID ABP51693 standard; peptide; 18 AA.
XX
AC ABP51693;
XX
DT 01-OCT-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:49.
XX
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200246238-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US047656.
XX
PR 05-DEC-2000; 2000US-0251448P.
PR

PR 04-MAY-2001; 2001US-0288889P.
PR 29-MAY-2001; 2001US-0294068P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Barbas-Frederickson S, Renshaw M;
XX
DR WPI; 2002-566610/60.
DR N-PSDB; ABQ73371.
XX
PT A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoetin mimetic.
XX
PS Claim 20; Fig 5; 113pp; English.
XX
CC The present invention describes an immunoglobulin molecule or its fragment
CC (i) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC promegakaryocytes or megakaryocytes, which results in increased platelet
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 5; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 2 TIEGPTLRQWLAAR 15
||:|||||||:|
||:|||||||:|

RESULT 18
ABP51691
ID ABP51691 standard; peptide; 18 AA.
XX
AC ABP51691;
XX
DT 01-OCT-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:45.
XX
KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
KW complementarity determining region; immunoglobulin; antianaemic;
KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200246238-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US047656.

XX
PR 05-DEC-2000; 2000US-0251448P.
PR 04-MAY-2001; 2001US-0288889P.
PR 29-MAY-2001; 2001US-0294068P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Barbas-Frederickson S, Renshaw M;
PI
DR WPI; 2002-566610/60.
DR N-PSDB; ABQ73369.
XX
PT A novel immunogen molecule comprising a region in which amino acid
PT residues corresponding to at least a portion of the complementary
PT determining region are replaced or fused with an erythropoietin or
PT thrombopoetin mimetic.
XX
PS Claim 20; Fig 5; 113pp; English.
XX
CC The present invention describes an immunoglobulin molecule or its fragment
CC (I) comprising a region where amino acid residues corresponding to at
CC least a portion of the complementary determining region (CDR) are
CC replaced or fused with biologically active peptides e.g. a peptide
CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
CC that is flanked with proline at its carboxy terminus. (I) has
CC antianaemic, haemostatic and nephrotropic activities, and can be used as
CC a stimulator of proliferation, differentiation and maturation of
CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
CC for stimulating proliferation, differentiation or growth of
CC promegakaryocytes or megakaryocytes, where (I) is contacted with
CC promegakaryocytes or megakaryocytes, which results in increased platelet
CC production. (I) with a region where amino acid residues corresponding to
CC a portion of CDR is replaced with an EPO mimetic, or which has one or
CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
CC production of red blood cells, where (I) is contacted with haematopoietic
CC stem cells or their progenitors. (I) is useful for diagnostics or
CC therapeutics, in cell isolation strategies, and for treating patients
CC suffering from deficiency in cell populations caused by disease,
CC disorders or treatments related to the suppression of haematopoiesis.
CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 5; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 2 TIEGPTLRQWLAAR 15
||:|||||||:|
||:|||||||:|

RESULT 19
ADQ16625
ID ADQ16625 standard; peptide; 18 AA.
XX
AC ADQ16625;
XX
DT 09-SEP-2004 (first entry)
XX
DE TPO mimetic peptide with random flanking residues SEQ ID NO:45.
KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.
XX
OS Unidentified.
XX
PN WO2004050017-A2.
XX
PD 17-JUN-2004.
XX

```
PF 17-NOV-2003; 2003WO-US036894.
XX
PR 02-DEC-2002; 2002US-00307724.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M;
XX
DR WPI; 2004-460973/43.
DR N-PSDB; ADQ16626.
XX
PT New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
PS Example 1; SEQ ID NO 45; 107pp; English.
XX
CC The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents a TPO mimetic peptide with flanking
CC residues.
XX
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 8; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db ||:|||||:|
2 TIEGPTLRQWLAAR 15

RESULT 20
ADQ16629
ID ADQ16629 standard; peptide; 18 AA.
XX
AC ADQ16629;
XX
DT 09-SEP-2004 (first entry)
XX
DE TPO mimetic peptide with random flanking residues SEQ ID NO:49.
XX
KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.
XX
OS Unidentified.
XX
PN WO2004050017-A2.
XX
PD 17-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-US036894.
XX
PR 02-DEC-2002; 2002US-00307724.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M;
XX
DR WPI; 2004-460973/43.
DR N-PSDB; ADQ16630.
XX
PT New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic

or a TPO mimetic, useful for treating thrombocytopenia.

Example 1; SEQ ID NO 49; 107pp; English.

The invention relates to a novel immunoglobulin molecule or its fragment
comprising a region where amino acid residues corresponding to at least a
portion of a two complementarity determining regions (CDRs) are replaced
with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
invention has immunosuppressive activity, and may have a use in
immunotherapy. The immunoglobulin molecule is useful for diagnosing or
treating thrombocytopenia as a result of chemotherapy, bone marrow
transplantation, or chronic diseases such as idiopathic thrombocytopenia.
The present sequence represents a TPO mimetic peptide with flanking
residues.

Sequence 18 AA;

Query Match 63.9%; Score 62; DB 8; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db ||:|||||:|
2 TIEGPTLRQWLAAR 15

RESULT 21
ADV44363
ID ADV44363 standard; peptide; 18 AA.
XX
AC ADV44363;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 45.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orenicia C;
XX
DR WPI; 2005-031588/03.
DR N-PSDB; ADV44364.
XX
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Example 1; SEQ ID NO 45; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
```

CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin molecule or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
||:||||||| :|
Db 2 TIEGPTLRQWLAAR 15

RESULT 22
ADV44367
ID ADV44367 standard; peptide; 18 AA.
XX
AC ADV44367;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 49.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-031588/03.
DR N-PSDB; ADV44368.
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Example 1; SEQ ID NO 49; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an

CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
Best Local Similarity 78.6%; Pred. No. 0.0081;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
||:||||||| :|
Db 2 TIEGPTLRQWLAAR 15

RESULT 23
AEB12840
ID AEB12840 standard; peptide; 18 AA.
XX
AC AEB12840;
XX
DT 08-SEP-2005 (first entry)
XX
DE TPO mimetic peptide, X7c.
XX
KW TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.
XX
OS Synthetic.
XX
PN WO2005060642-A2.
XX
PD 07-JUL-2005.
XX
PF 15-DEC-2004; 2004WO-US041946.
XX
PR 15-DEC-2003; 2003US-00737290.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-479402/48.
DR N-PSDB; AEB12841.
XX
PT New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
PS Example 1; SEQ ID NO 49; 152pp; English.
XX
CC The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under

DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 19 AA;
Query Match 63.9%; Score 62; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0086;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSREHTS 18
Db 5 GPTLRQWLAAARTHLS 19
RESULT 26
AAW36644
ID AAW36644 standard; peptide; 19 AA.
XX
AC AAW36644;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX

PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 26; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 19 AA;
Query Match 63.9%; Score 62; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0086;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSREHTS 18
Db 5 GPTLRQWLAAARTHLS 19
RESULT 27
AAU25863
ID AAU25863 standard; peptide; 19 AA.
XX
AC AAU25863;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #49.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

XX PF 01-MAR-2000; 2000US-00516704.
XX PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
XX DR
XX PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX PS
PS Disclosure; Col 147; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R) . Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk.The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 14 AA;
Query Match 59.8%; Score 58; DB 4; Length 14;
Best Local Similarity 71.4%; Pred. No. 0.027;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 IKGPTLRQWLKSPRE 15
Db :|||||:|:
1 IEGPTLRQWLAAARK 14
RESULT 30
AAW09460
ID AAW09460 standard; protein; 18 AA.
XX
AC AAW09460;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding compound peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .18
FT /note= "Preferably linkages are selected from: -
FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
FT lower alkyl"

FT Modified-site 1 /note= "Preferably N-terminus is selected from: -NRR1; -
FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
FT benzyloxycarbonyl-NH; benzyloxycarbonyl-NH with 1-3
FT substitutions on the phenyl ring selected from lower
FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
FT independently selected from hydrogen and lower alkyl"
FT 18
FT Modified-site /note= "Preferably C-terminus is -C(O)R2 where R2 is
FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
FT and R4 are independently selected from hydrogen and lower
FT alkyl, and where the nitrogen atom of the -NR3R4 group
FT can optionally be the amine group of the N-terminus of
FT the peptide forming a cyclic peptide"
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
DR
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 18; Page 89; 106pp; English.
XX
CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding
CC affinity to TR as expressed by an IC50 of no more than about 100 mum. The
CC compound (especially if modified, see features table) can be used for
CC treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX
SQ Sequence 18 AA;
Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSR 14
Db :|||||:|:
1 SIEGPTLREWLTSR 14
RESULT 31
AAW09498
ID AAW09498 standard; protein; 18 AA.
XX
AC AAW09498;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;

KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
XX 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX) GLAXO GROUP LTD.
PA
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 27; 106pp; English.
XX
XX The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db :|:|||||:|
1 SIEGPTLREWLTSR 14

RESULT 32
AAW36649
ID AAW36649 standard; peptide; 18 AA.
XX
AC AAW36649;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX WO9640750-A1.
PN
XX 19-DEC-1996.
PD
XX
PF 07-JUN-1996; 96WO-US009623.
PR
XX
XX Thrombopoietin receptor binding peptide.
DE
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX WO9640750-A1.
PN
XX 19-DEC-1996.
PD
XX
PF 07-JUN-1996; 96WO-US009623.
PR
XX
XX 07-JUN-1995; 95US-00478128.
PR

PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 27; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db :|:|||||:|
1 SIEGPTLREWLTSR 14

RESULT 33
AAW33027
ID AAW33027 standard; peptide; 18 AA.
XX
AC AAW33027;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX WO9640750-A1.
PN
XX 19-DEC-1996.
PD
XX
PF 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00485301.
PR
XX (GLAX) GLAXO GROUP LTD.
PA
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
DR
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
PS Claim 19; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopaenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLREWLTSR 14
:|:|||||:|

RESULT 34
AAW36652
ID AAW36652 standard; peptide; 18 AA.
XX
AC AAW36652;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
DR WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 27; 106pp; English.
XX
CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLREWLTSR 14
:|:|||||:|

RESULT 35
AAB17026
ID AAB17026 standard; peptide; 18 AA.
XX
AC AAB17026;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:82.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 222; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 3; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

		Matches	10;	Conservative	3;	Mismatches	1;	Indels	0;	Gaps	0;								
QY	1	TIKGPTLRQWLKSR	14																
		: : : :																	
Db	1	SIEGPTLREWLTSR	14																
RESULT 36																			
AAU25868																			
ID	AAU25868	standard; peptide; 18 AA.																	
XX	AC	AAU25868;																	
XX	DT	17-DEC-2001 (first entry)																	
XX	DE	Human thrombopoietin receptor (TPO-R) activator peptide #54.																	
XX	KW	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;																	
KW	KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;																	
KW	KW	bone marrow transplantation; haematological disorder; platelet disorder;																	
KW	KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;																	
KW	KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;																	
KW	KW	in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.																	
XX	OS	Homo sapiens.																	
XX	PN	US6251864-B1.																	
XX	PD	26-JUN-2001.																	
XX	PF	01-MAR-2000; 2000US-00516704.																	
XX	PR	07-JUN-1995; 95US-00478128.																	
PR	PR	07-JUN-1995; 95US-00485301.																	
PR	PR	07-JUN-1996; 96WO-US009623.																	
PR	PR	15-AUG-1996; 96US-00699027.																	
XX	PA	(GLAX) GLAXO GROUP LTD.																	
XX	PI	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;																	
PI	PI	Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;																	
PI	PI	Yin Q;																	
XX	XX	WPI; 2001-564142/63.																	
DR	DR	Activating thrombopoietin receptors in cells, used to treat																	
XX	XX	thrombocytopenia and hematological disorders, comprises contacting cells																	
PT	PT	with peptides and peptide mimetics attached to hydrophilic polymers.																	
XX	XX	Disclosure; Col 20; 128pp; English.																	
PS	PS	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that																	
XX	XX	bind to and activate the human thrombopoietin receptor (TPO-R). Methods																	
CC	CC	of activating thrombopoietin receptors in cells comprise contacting the																	
CC	CC	cells with effective amounts of peptides and peptide mimetics attached to																	
CC	CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such																	
CC	CC	as that due to chemotherapy, radiation therapy or bone-marrow																	
CC	CC	transplantation and to prevent thrombocytopenia in patients at risk.The																	
CC	CC	sequences are used to treat and prevent haematological disorders																	
CC	CC	including thrombocytopenia and platelet disorders. They are used in vitro																	
CC	CC	as unique tools for understanding the biological role of thrombopoietin																	
CC	CC	(TPO) and to develop other compounds that bind to and activate the TPO																	
CC	CC	receptor. The peptides can be used to detect TPO receptors on living																	
CC	CC	cells and fixed cells, in biological fluids, in tissue homogenates, and																	
CC	CC	in purified or natural biological materials. They may also be used for in																	
CC	CC	situ staining, fluorescence-activated cell sorting, Western blotting and																	
CC	CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can																	
CC	CC	be used for in vitro expansion of megakaryocytes and their committed																	
CC	CC	progenitors alone or in conjunction with additional cytokines																	
XX	XX	Sequence 18 AA;																	
SQ	SQ	Query Match 59.8%; Score 58; DB 4; Length 18;																	
		Best Local Similarity 71.4%; Pred. No. 0.035;																	

Best Local Similarity 71.4%; Pred. No. 0.035; Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
:|:|||||:|
Db 1 SIEGPTLREWLTSR 14

RESULT 38
AAU25871
ID AAU25871 standard; peptide; 18 AA.
XX AAU25871;
XX
DT 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #57.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.
OS
XX US6251864-B1.
PN
XX
PD 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.
PF
XX
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00485301.
PR
XX 07-JUN-1996; 96WO-US009623.
PR
XX 15-AUG-1996; 96US-00699027.
XX
XX
PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.

XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 20; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 4; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
:|:|||||:|
Db 1 SIEGPTLREWLTSR 14

RESULT 39
ABB72912
ID ABB72912 standard; peptide; 18 AA.
XX
AC ABB72912;
XX

DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:82.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.
OS
OS Synthetic.

XX WO200183525-A2.
PN
XX
PD 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 44; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 5; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLREWLTSR 14
:|:|||||:|

RESULT 40
ADJ73064
ID ADJ73064 standard; peptide; 18 AA.
XX
AC ADJ73064;
XX
DT 06-MAY-2004 (first entry)
XX TPO mimetic peptide sequence SeqID 518.
DE
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 518; 97pp; English.
XX

This invention relates to novel mammalian CDR mimetibodies, specific
portions or variants thereof. Specifically, it refers to an antibody
fragment where a protein has been inserted into, or replaces a portion
of, one or more CDR regions, such that each CDR mimetibody comprises at
least one portion of a heavy chain or light chain variable region, which
itself comprises at least one human framework region and at least one
ligand binding region (LBR). The present invention describes human
mimetibodies, including modified immunoglobulins and cleavage products
that can be useful in gene therapy and the generation of transgenic
plants and animals. Furthermore, the CDR mimetibody is useful for
preparing compositions for modulating, treating or reducing the symptoms
of immune, cardiovascular, infectious, malignant and/ or neurologic
diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
cardiant, antimicrobial, cytostatic and neuroprotective activities. This
peptide sequence is a TPO mimetic peptide sequence used to make a
mimetibody of the invention.

Sequence 18 AA;
XX
SQ

Query Match 59.8%; Score 58; DB 7; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLREWLTSR 14
:|:|||||:|

RESULT 41
ADJ52699
ID ADJ52699 standard; peptide; 18 AA.
XX
AC ADJ52699;
XX
DT 06-MAY-2004 (first entry)
XX
DE
XX CHI deleted mimetibody-related peptide SeqID518.
KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 518; 129pp; English.
XX

This invention relates to CHI deleted mimetibodies (and the DNA sequences
which encode them), compositions, methods and uses. The invention may be
useful for the development of compounds with an immunosuppressive,
cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
antibacterial, virucide or fungicide activity. In addition, the disclosed
sequences may prove useful for gene therapy. The CHI-deleted mimetibody
is useful for diagnosing or treating a disease condition in a cell,
tissue, organ or animal, specifically for modulating, treating,
alleviating, preventing the incidence or reducing the symptoms of an
immune, cardiovascular (for example arrhythmia, hypertension or heart
failure), or neurodegenerative (for example multiple sclerosis, dementia
or Alzheimer's disease) diseases or disorders, anaemia, cancerous
conditions, or infectious diseases (for example bacterial, viral or
fungal infection). The present sequence is that of a peptide which may be
used during the creation of a mimetibody of the invention.

Sequence 18 AA;
SQ

Query Match 59.8%; Score 58; DB 8; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
:|:|||||:|
Db 1 SIEGPTLREWLTSR 14

RESULT 42
ADJ51660
ID ADJ51660 standard; peptide; 18 AA.
XX
AC ADJ51660;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID518.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 518; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurologic, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 18 AA;
Query Match 59.8%; Score 58; DB 8; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSR 14
:|:|||||:|
Db 1 SIEGPTLREWLTSR 14

RESULT 43
ADV44468
ID ADV44468 standard; peptide; 18 AA.
XX
AC ADV44468;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 152.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orenacia C;
XX
DR WPI; 2005-031588/03.
XX
PS New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Example 8; SEQ ID NO 152; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to

CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.035;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRHT 17
Db 3 IEGPTLRQWLAAARANS 18
|:|||||:|:|:
3 IEGPTLRQWLAAARANS 18

RESULT 44
AEB12939
ID AEB12939 standard; peptide; 18 AA.
XX
AC AEB12939;
XX
DT 08-SEP-2005 (first entry)
XX
DE TPO mimetic peptide for VH CDR2 grafting.
XX
KW TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.
XX
OS Synthetic.
XX
PN WO2005060642-A2.
XX
PD 07-JUL-2005.
XX
PF 15-DEC-2004; 2004WO-US041946.
XX
PR 15-DEC-2003; 2003US-00737290.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-479402/48.
XX
PT New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
PS Example 8; SEQ ID NO 152; 152pp; English.
XX
CC The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under
CC conditions suitable for expression of the immunoglobulin or its
CC fragment), a composition comprising an immunoglobulin (or its fragment)
CC and a pharmaceutically acceptable carrier, treating congestive heart
CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
CC glucagons, glucagon mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin

CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1),
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to
CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a Thrombopoietin (TPO) mimetic , erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.035;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRHT 17
Db 3 IEGPTLRQWLAAARANS 18
|:|||||:|:|:
3 IEGPTLRQWLAAARANS 18

RESULT 45
ADQ16705
ID ADQ16705 standard; protein; 128 AA.
XX
AC ADQ16705;
XX
DT 09-SEP-2004 (first entry)
XX
DE Modified immunoglobulin clone 116 HC variable region SEQ ID NO:125.
XX
KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.
XX
OS Synthetic.
XX
PN WO2004050017-A2.
XX
PD 17-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-US036894.
XX
PR 02-DEC-2002; 2002US-00307724.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M;
XX
DR WPI; 2004-460973/43.
XX
PT New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.
XX
PS Claim 9; SEQ ID NO 125; 107pp; English.
XX
CC The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced

CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents immunoglobulin clone 116 heavy chain
CC variable region.
XX
SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 8; Length 128;
Best Local Similarity 62.5%; Pred. No. 0.34;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHT 17
|:||||| :| :|
Db 52 IEGPTLRQWLAARANS 67

RESULT 46
ADV44466
ID ADV44466 standard; protein; 128 AA.
XX
AC ADV44466;
XX
DT 10-MAR-2005 (first entry)
XX
DE Anti-tetanus toxoid Fab pAX116 heavy chain variant PRL5-116 13F2.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; pAX116.
XX
OS Homo sapiens.
OS Synthetic.
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-031588/03.
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Claim 148; SEQ ID NO 150; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;

CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of a variant of the anti-tetanus toxoid
CC antibody pAX116 heavy chain, CDR2 of which has been partially replaced
CC with residues of a TPO-mimetic peptide.
XX
SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 9; Length 128;
Best Local Similarity 62.5%; Pred. No. 0.34;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHT 17
|:||||| :| :|
Db 52 IEGPTLRQWLAARANS 67

RESULT 47
ADV44463
ID ADV44463 standard; protein; 128 AA.
XX
AC ADV44463;
XX
DT 10-MAR-2005 (first entry)
XX
DE Anti-tetanus toxoid Fab pAX116 heavy chain variant PRL5-116.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; pAX116.
XX
OS Homo sapiens.
OS Synthetic.
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orencia C;
XX
DR WPI; 2005-031588/03.
XX
PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Example 11; SEQ ID NO 147; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also

disclosed are: engineering immunoglobulin molecules or fragments; creation of a library of monoclonal antibodies; stimulating proliferation, differentiation, or growth of megakaryocytes; and activating a homodimeric receptor protein. The immunoglobulin molecules are useful for treating HIV-infected patients, patients undergoing chemotherapy, bone marrow transplant patients, stem cell transplant patients, or patients suffering from myeloproliferative disorders. This is the amino acid sequence of a variant of the anti-tetanus toxoid antibody PAX116 heavy chain, CDR2 of which has been partially replaced with residues of a TPO-mimetic peptide.

Sequence 128 AA;

immunoglobulin or its fragment and a pharmaceutical carrier. Also disclosed are: engineering immunoglobulin molecules or fragments; creation of a library of monoclonal antibodies; stimulating proliferation, differentiation, or growth of megakaryocytes; and activating a homodimeric receptor protein. The immunoglobulin molecules are useful for treating HIV-infected patients, patients undergoing chemotherapy, bone marrow transplant patients, stem cell transplant patients, or patients suffering from myeloproliferative disorders. This is the amino acid sequence of a variant of the anti-tetanus toxoid antibody PAX116 heavy chain, CDR2 of which has been partially replaced with residues of a TPO-mimetic peptide.

Sequence 128 AA;

Claim 148; SEQ ID NO 151; 139pp; English.

The invention describes an immunoglobulin molecule or its fragment comprising: a region where amino acid residues corresponding to at least a portion of two CDRs are replaced with a peptide mimetic selected from an EPO mimetic or a TPO mimetic; or a region where amino acid residues corresponding to at least a portion of a CDR is replaced by a peptide mimetic including SEQ ID NO. 126 (not defined in the specification), where X at each occurrence represents any amino acid. Also described are: a nucleic acid encoding an immunoglobulin molecule or its fragment; an expression vector comprising the nucleic acid of (1); a host cell transformed with the expression vector of (2); producing an immunoglobulin molecule or its fragment; and a composition comprising the

CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of a variant of the anti-tetanus toxoid
CC antibody pAX116 heavy chain, CDR2 of which has been partially replaced
CC with residues of a TPO-mimetic peptide.
XX
SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 9; Length 128;
Best Local Similarity 62.5%; Pred. No. 0.34;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
|:|||||:|:|:
Db 52 IEGPTLRQWLAAARANS 67

RESULT 50
ADV44441
ID ADV44441 standard; protein; 128 AA.

XX AC ADV44441;

XX DT 10-MAR-2005 (first entry)

XX DE pAX116 variable heavy chain variable region.

XX KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; pAX116.

XX OS Homo sapiens.

XX OS Synthetic.

XX OS Unidentified.

XX PN WO2004108078-A2.

XX PD 16-DEC-2004.

XX PF 26-MAY-2004; 2004WO-US016574.

XX PR 02-JUN-2003; 2003US-00452590.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Frederickson S, Renshaw M, Orenca C;

XX DR WPI; 2005-031588/03.

XX PT New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.

XX PS Example 8; SEQ ID NO 125; 139pp; English.

XX CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell

CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of the variable region of modified anti-
CC tetanus toxoid antibody pAX116 heavy chain that comprises two agonist TPO
CC -mimetic peptides.
XX
SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 9; Length 128;
Best Local Similarity 62.5%; Pred. No. 0.34;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
|:|||||:|:|:
Db 52 IEGPTLRQWLAAARANS 67

Search completed: May 12, 2006, 10:35:37
Job time : 143.869 secs

10/632,388